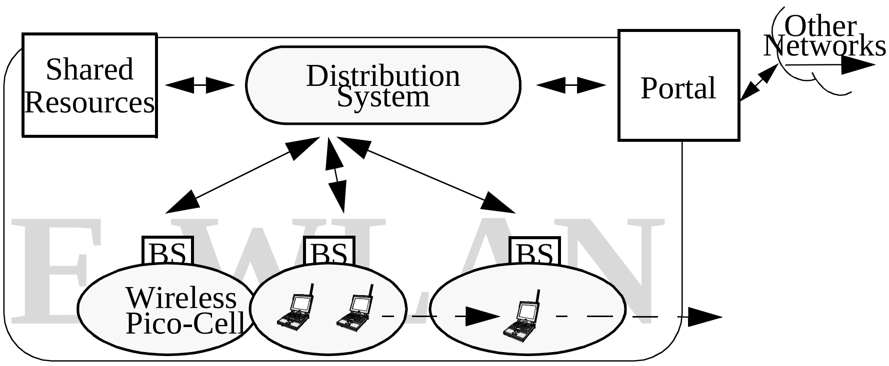
Grouping Wireless Picocells to build a Local Area Wireless Infrastructure

Fast and efficient Local Mobility Management and Range Extension

Dipl.-Ing. Jost Weinmiller



A Thesis submitted in partial Fulfillment of the Requirements for the
Degree of Doctor of Engineering (Dr.-Ing.)
at the Technical University of Berlin

July 2000
1st full Version: January 1998
2nd (revised) Version: January 1999

D83

Grouping Wireless Picocells to build a Local Area Wireless Infrastructure

Fast and efficient Local Mobility Management and Range Extension

vorgelegt von
Diplom-Ingenieur
Jost Weinmiller

Vom Fachbereich 12 - Elektrotechnik
der Technischen Uniersität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurswissenschaften
- Dr.-Ing. -
genehmigte Dissertation

Promotionsausschuß:
Vorsitzender: Prof. Dr.-Ing. Filbert
Berichter: Prof. Dr.-Ing. Wolisz
Berichter: Prof. Dr.-Ing. David

Tag der wissenschaftlichen Aussprache: 5. April 2000

Berlin 2000
D 83

Grouping Wireless Picocells to build a Local Area Wireless Infrastructure

Fast and efficient Local Mobility Management and Range Extension

Dipl.-Ing. Jost Weinmiller

Abstract

The thesis presents a systematic investigation of solutions to provide efficient handover in a wireless- based LAN-type communication infrastructure in an in-door environment. The chosen architecture builds upon several wireless islands and interconnects them via a second networking infrastructure of a different type. Through this several cells are grouped together into a larger administrative unit of the dimensions of a common LAN. This architecture is advised since the coverage area of future cells will be very small (~10 meter) causing very frequent handover. In order to avoid the "expensive" handover on the network layer at such high frequency a single network with multiple cells is built up, allowing for a local area mobility management that can be more efficient and faster than network layer based solutions.

With respect to the increasing demand for support for traffic with real-time requirements in communication networks, the capabilities and the issues arising with real-time communication in the system in question are furthermore discussed. The thesis investigates the space for applicable strategies and identifies distinctive basic solutions. Specifications for those solutions are developed. The specified approaches have been compared and evaluated regarding their performance and their different functional properties, with the help of simulation. The thesis finishes with conclusive remarks.

Gruppieren von drahlosen Picozellen um eine lokale drahtlose Infrastruktur aufzubauen

Schnelles und effizientes lokales Mobilitätsmanagement

Dipl.-Ing. Jost Weinmiller

Zusammenfassung

Diese Dissertation stellt eine systematische Untersuchung vor von Lösungen für effizienten Hand-over in einer drahtlosen, LAN-basierten Kommunikationsumgebung in einer geschlossenen Umgebung. Die gewählte Umgebung besteht aus mehreren drahtlosen Zellen die mittels einer zweiten Netzwerkinfrastruktur verbunden werden Die Funkzellen werden dadurch zu einer größeren Einheit zusammengebunden, die den gängigen Dimensionen eines lokalen Netzes entsprechen. Diese Architektur wurde gewählt da die Reichweite zukünftiger Funkzellen sehr klein sein wird (~10 Meter) was zu sehr häufigem Handover führen wird. Um die aufwändigen Handover auf der Netzwerkschicht bei dieser hohen Rate zu vermeiden wird ein einzelnes lokales Netz bestehend aus mehreren Zellen aufgebaut, was ein Mobilitätsmanagement im lokalen Rahmen erlaubt und das dadurch effizienter und schneller ist als Lösungen auf Netzwerkschicht.

Hinsichtlich der steigenden Nachfrage nach Unterstützung für Echtzeitkommunikation in Kommunikationsnetzen, werden die offenen Fragen und diesbezüglichen Eigenschaften in solch einem System weitergehend diskutiert und betrachtet. Die Dissertation untersucht den Raum für hier anwendbare Strategien und identifiziert die Grundmechanismen, die ihnen zugrunde liegen. Für die Lösungen werden Spezifikationen erarbeitet. Die spezifizierten Schemata wurden mittels simulativer Verfahren bezüglich ihrer Leistung verglichen und bewertet im Hinblick auf ihre unterschiedlichen Funktionalitäten. Die Dissertation schliesst mit abschliessenden Bemerkungen ab.

Acknowledgments

First of all, I would like take the opportunity to express my strongest discomfort with the system-immanent weak position of the doctoral student in the process of promotion, leaving him in almost unlimited dependency to his advisor. This permits the latter to abuse his position in many ways, e.g. to pursue rather his own interest or by disrespectfully neglecting to advise over undue long timespans, while the doctoral student is left with no procedural means to ensure fair and timely treatment. This structural dilemma cannot be justified to aid good scientific education nor does it help the researcher to achieve scientific excellence.

Nevertheless I want to thank my advisor, Professor Adam Wolisz, who provided valuable input, guidance and motivation in the course of the work. The very enjoyable communicative and creative atmosphere, that he encourages among the members of the research group has proven to be most beneficial for this work.

I especially would like to thank Andreas Festag, who did very important work on the multicast protocol as part of his master thesis. His input, his major contribution in setting up the fundament for the simulation model and - most importantly - his reliability as a friend were indispensable for the success of this thesis.

Thanks are also due to all the members of the TKN-group, for their co-operation and help that made my life easier and enjoyable, especially mentioning my good friends Hagen Woesner and Frank Fitzek. Thank goes to Torsten Koehler in Stockholm for his help in proof reading.

Special thank goes to the Deutsche Forschungsgemeinschaft (DFG) for providing financial support within the priority research funding program ‘Mobile Communication’.

Table of Contents

List of Figures

Chapter 1

Introduction

1. Introduction

Wireless local area networks (WLANs) for indoor-use are currently one of the fastest growing market segments in the data networking market. First products, operating in the license free ISM frequency band at 900 MHz, appeared on the market a few years ago. They offer data rates below 1 Mbit/s with proprietary architectures and interfaces lacking the possibility for interoperability. The main additional feature they offered, compared to wired communication interfaces, was the liberation of the end system from the cord, thus allowing for fast and easy installation and reconfiguration. However none of the products provides support for **communication on the move**, i.e. if the host with the wireless connection wants to leave the range of the current coverage area of the cell, manual reconfiguration of the end system and/or the base station is required, or - if automatic procedures are available - interruptions in the communication of unspecified, possibly very long duration occur.

For the second generation of WLANs many vendors and researchers recognized the need for interoperability and that this can only be achieved through standardization (comparable to the development of wired LAN communication where standardization of Ethernet and Token Ring was the decisive factor for the explosive development of the market). The Institute of Electrical and Electronics Engineers (IEEE) founded a working group within the IEEE-802 framework to standardize WLAN systems - the now complete IEEE 802.11 standard for “Wireless LAN Medium Access Control and Physical Layer Specification” [3]. Its operating frequency is 2.4 GHz[1], data rates of 1 Mbit/s and 2 Mbit/s can be used[2]. The coverage area of the (just recently shipped) products will be some tens of meters, probably a couple of office rooms since some wall penetration can be expected. Similarly the European Telecommunications Standards Institute (ETSI) developed a standard for a High PERformance wireless LAN (HIPERLAN)[4], that is operating at 5.4 GHz.

1.1. Small Picocells for Wireless In-door LANs

Already in progress is work on the third generation of wireless LANs. These systems will have an even smaller coverage area than its preceding generations for a number of reasons:

- The application of higher frequency ranges (5GHz, 17 GHz, 60 GHz) - necessary to achieve higher data rates - results in a smaller propagation radius and worth wall penetration. This is caused by the inverse dependency of received power and operation frequency[3]:

$$P_r = P_t \cdot \left(\frac{c}{4\pi df}\right)^2$$

P_r	received power
P_t	sending power
c	light speed
d	distance
f	frequency

1. 3 different physical layers are defined: Baseband Infrared (BBIR), Direct Sequence Spread Spectrum (DSSS), Frequency Hopping Spread Spectrum (FHSS). The latter two operate in the 2.4 GHz frequency range.

2. Currently extensions to IEEE 802.11 are being developed to provide 10MB/s at an operating frequency of 5GHz

3. This equation is valid for propagation through vacuum. For other propagation mediums (walls, air, ...) the equation has to be extended by a constant factor reflecting the propagation properties of the medium.

- Another possibility to get a larger share of the available bandwidth for each station is to reduce the number of stations in each cell. This also improves the performance of the multiple access protocols that will have to operate in overload condition much less frequently. However in order to accommodate the same number of stations per area smaller cells are needed.

- Since battery power is still a very scarce resource in mobile electronic devices it is desirable to transmit with lower transmission power to save battery (this also helps in order to reduce the electromagnetic smog caused by radiation emitting devices). The propagation radius of course gets smaller respectively.

- The smaller the diameter of a cell is the closer the frequency can be reused. Therefore one gets a higher frequency reuse factor and correspondingly a higher spectrum efficiency with smaller cells

For these reasons one can assume that future WLAN cells will have diameters of just several meters while covering maximal one room.

1.2. Problems

An organization deciding to build up an infrastructure consisting of such small wireless communication cells will face two problems: first it will only get a number of unconnected wireless islands each covering a single-room-size cell. These communication islands will have to be integrated into other existing local wired infrastructures, i.e. other local area networks installed parallel or to the internetwork architectures connecting to outside networks.

The second problem that has to be solved is the accommodation of mobile users in the networking architecture. Wireless communication technology allows the user to cut the cord that bonds the end system to a fixed location. Users will want to use this newly gained freedom to communicate while being on the move. Since currently used network architectures are designed to accommodate only static or semi-static end systems one cannot assume to automatically get a proper communication environment for mobile users. This can for example be seen in the addressing structure applied in the internet. Figure 1 shows some real network addresses and their location. On a sidenote it is interesting to see that the numerical addresses do not provide any information on the location of the network. However the fact that the networks and thus the hosts located in the network are specified by an address that is used as location determination **and** name at the same time, makes it impossible for a host to move into another network while maintaining its identity - its name.

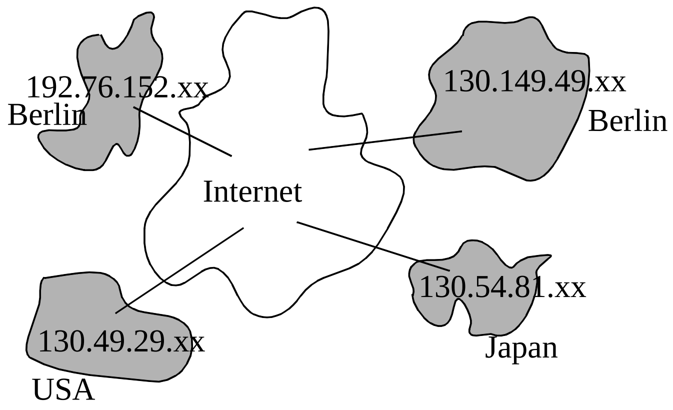


Figure 1:Names and Locations in the Addressing Structure of the Internet Architecture

Once a mobile end system moves out of its current cell and moves into another one the underlying infrastructure has to deal with handover management, i.e. dealing with currently active data communication and with location management, i.e. getting the updated location information to the place where it is needed.

1.3. Mobility

Regarding the type of solution offered to mobile end systems one may choose to implement better, more complex solutions or less featured but simpler solutions for the mobility management.

Figure 2 shows the problem space for mobile communication. One may want to allow only users that move around freely but only communicate while being stationary. Maybe one may want to let users communicate while moving at low speed (e.g. walking), or even at high speed (e.g. car speed, which will not be a frequent case in an in-door environment). One may want to offer users a limited set of functionality while communicating mobile (e.g. only POP-based email service) or will attempt to keep the full fledged set of services (e.g. FTP, WWW, NFS, voice/video transmission, conferencing, whiteboard,...) with all their distinct features even in the event of mobility. Lastly one may want to settle for mobility management that involves manual interaction into the systems configuration, or prefer to get fully automatic mobility management that is transparent to the user.

The chosen in-house environment naturally limits the speed at which one will expect users to move therefore allowing for slow moving terminals is likely to be sufficient. Of course if manual intervention is required in-between each cell change the solution will unnecessarily slow down the possible speed at which the mobile user can move around therefore fully automatic mobility management is required. And of course the user will not agree to give up much of the functionality that he is used to get from the wired counterparts. To sum up: the communication system designed for this environment has to be designed to operate fully automatic while maintaining high-level functionality for systems moving at medium (i.e. walking) speed within the range of a typical local area network.

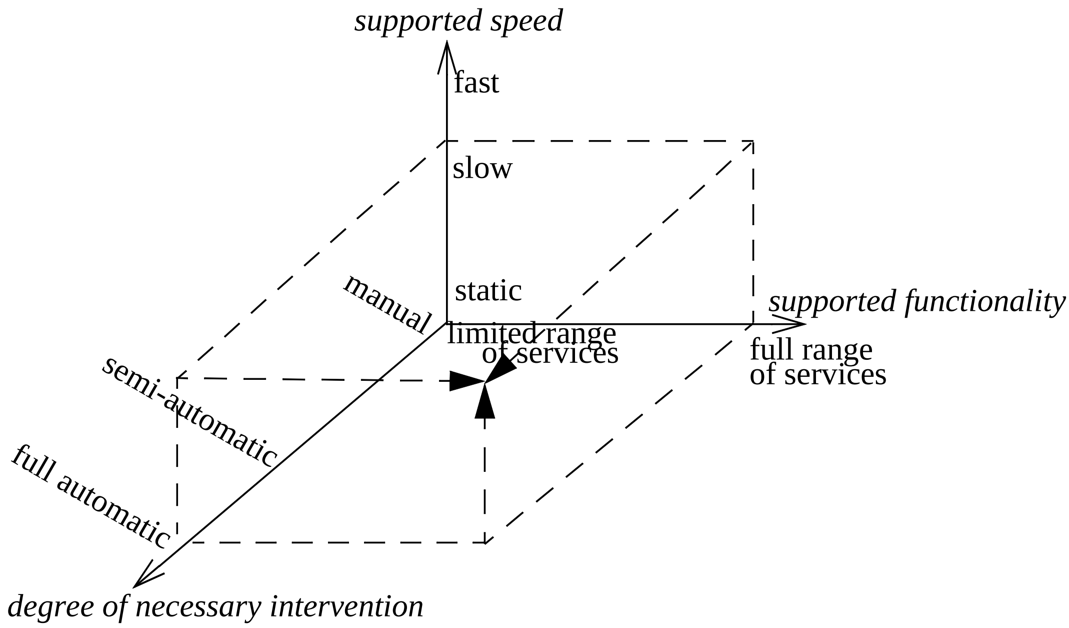


Figure 2:Problem Space of Mobile Communication

1.4. Grouping Wireless LAN Cells to Form a Larger Architecture

As discussed above, the coverage provided by a single pico-cell will definitely not be sufficient for a communication environment with the dimensions of a commonly used LAN in terms of number of supported stations and coverage area. Therefore one may want to extend the range of the communication area beyond the limits of a single pico cell. In order to also use the possibility for mobile computing offered inherently by the wireless medium one also needs to implement mobility management.

One approach to achieve the desired extension of the coverage area is to connect a group of pico-cells by a wired backbone. Such an architecture is shown in (Figure 3). The communication architecture consists of *several* wireless cells with limited diameter. Each cell has a base station (BS). All BSs are interconnected over a communication medium - the distribution medium (DM), which may be wire-based and thus will offer higher bit rates and significantly lower error rates than the wireless-based transmission. The wireless LAN is connected to other networks by a Portal. Shared resources like printer or file server can be connected directly to the distribution system (DS). The DS consists of the distribution medium, the associated interfaces in the base stations, portal and shared resources, and the functionality that implement the services of the distribution system. All together these elements form an extended wireless LAN, which will be called an E-WLAN from now on to make it distinctive from WLANs.

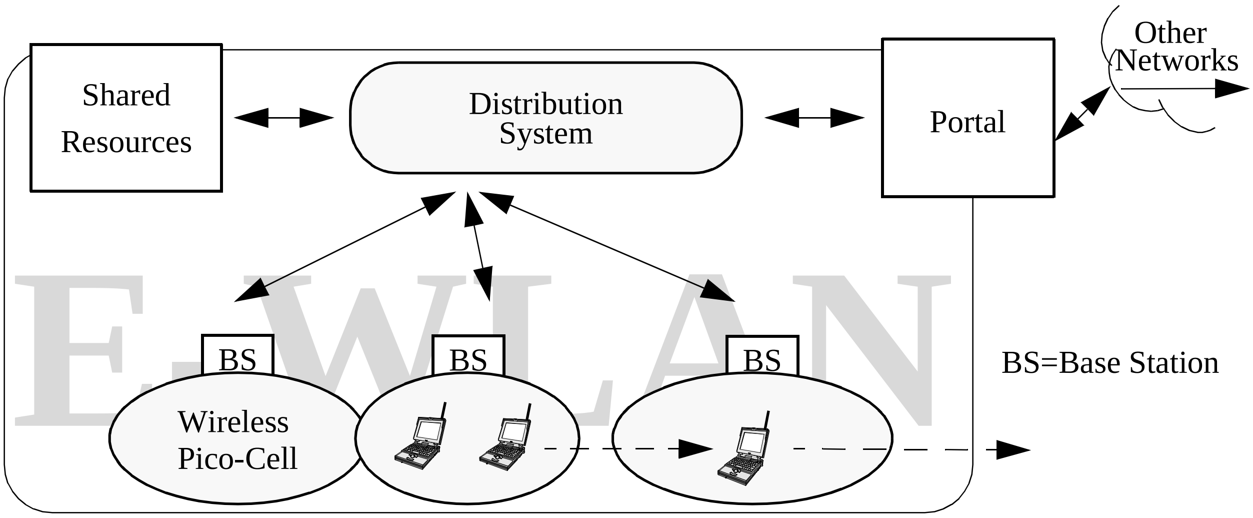


Figure 3:Architecture for Extended Wireless LAN (E-WLAN)

A mobile host (MH) may communicate while located within any of the participating cells. During a transition from one cell to another a handover occurs. The wireless cells may overlap or may be disjoint. In the process of a handover packets may be lost or duplicated, interruptions of connectivity of a certain length will occur. If the cells do not overlap the period of time that the mobile host spends in this uncovered area will simply add to this interruption time. The management of a handover obviously has significant influence on the experienced quality of service, especially in the case of high mobility, high bitrates and delay-sensitive data.

The architecture is oriented at the architecture chosen for the IEEE 802.11 standard[3], that similarly applies multiple “basic service sets” that each consist of one cell, that are interconnected over a “distribution system” into an “extended service set”. However the IEEE 802.11 standard focuses on the over-the-air interactions of wireless LAN products and does not provide a complete specification for a wireless LAN system. The internals of the distribution system are not defined by IEEE 802.11 and there is no exposed interface to the distribution system.

The IEEE 802.11 standard describes two operating modes for WLANs - infrastructure-based WLANs and ad-hoc-WLANs. Quite naturally, with the application of a backbone infrastructure - the distribution medium - this thesis only considers the infrastructure mode.

Approaches to provide support for terminal mobility are usually operated on the network layer (e.g. Mobile IP) or by the transport layer (e.g. ITCP), involving multiple separate networks. However within a single administrative domain of an E-WLAN, as defined above, support for terminal mobility within the data link layer seems to be very attractive. It allows for local mobility management rather than global mobility management. Intra-cell mobility is then divided into intra-LAN mobility (local mobility between different wireless cells of one E-WLAN) and inter-LAN mobility (mobility between wireless cells of two different E-WLANs). It can be assumed the most cell changes will occur within the local E-WLAN - users will only rarely move into another domain, therefore such local mobility management seems like the natural thing to do. This thesis will focus on the presentation of this alternative. Figure 4 shows the protocol stack in the base station and the mobile host that results from this architecture. The syntax of the TCP (or other transport layer) protocol as well as the IP syntax remains unchanged. Both protocols should remain unaffected by the underlying additional functionality, unless it is desired to provide inter-layer information e.g. indication that the interruption is handover related and not a result from congestion.

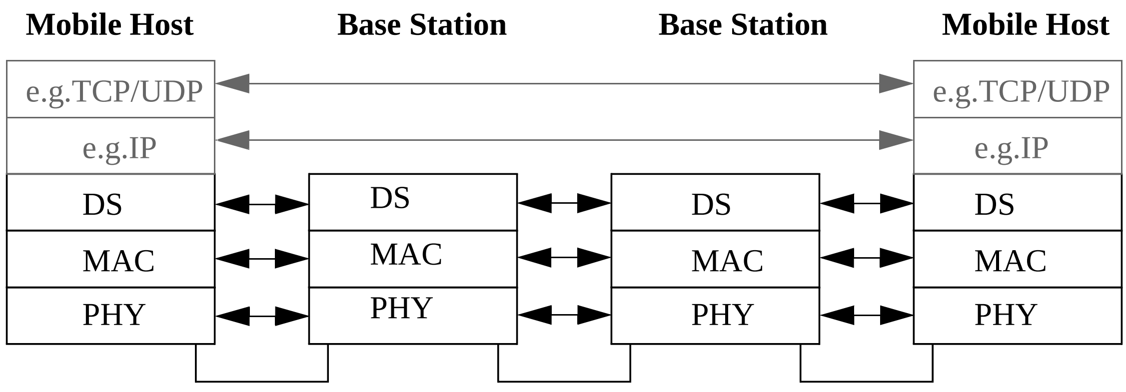


Figure 4:Protocol Stack for E-WLAN with Distribution System

1.5. Issues

The distribution system must enable transparent grouping of cells to form a **single E-WLAN**. For hosts outside of the E-WLAN the separation of the E-WLAN into its cells and distribution system is fully transparent, but also the mobile hosts (sender and receiver) should be kept as unaware as possible of the architectural structure. This allows to minimize the active participations of the mobile hosts with its limited supply of power supply and its connection to the wireless medium with its small bandwidth. A mobile host must be able to move from one cell to another (handover) without experiencing significant decrease of the quality of the transmission. An almost-continuous and relatively reliable network connection in spite of possibly frequent handover shall be offered. The distribution system has to provide unconstrained support for mobility among the connected pico-cells. To do so it has to have the means to track down the current location of a mobile host and has to deal with packets addressed to the old, no longer current location after a handover.

The IEEE 802.11 standard defines, that the distribution system has to provide services for

- Authentication and Privacy
- Association, Re- and Disassociation
- Distribution and Integration

The services Authentication and privacy are outside of the scope of this thesis. The solutions applied for these issues are independent from the solutions for last service. The services of association, re-association and disassociation are used but not otherwise discussed. This thesis will discuss and provide solutions for the problem of distribution and integration.

1.6. Application Environment

The application environment that the system will be designed for, consists of an in-door, office-type setting, as it is common for the wide spread wired local area networks. The setting limits the possible mobility to walking speed, at maximum *bicycle speed* (~20 km/h), the office environment excludes speeds above this value. Communication characteristics will be the same as for the wired LANs, since the same users will be using the networking infrastructure for the same purposes. Traffic will equally be split up (50/50 ratio) between traffic remaining within the local LAN and traffic going to external hosts. All traffic coming from or going to external destinations is usually sent through one designated host (gateway). Therefore every mobility related changes in the end-to-end path remain within the area of the local domain for local moves. The portion of the path beyond the gateway remains unchanged. This fact obviously gives advantage to an architecture that does not involve any host in the mobility management that is located beyond the local gateway outside the local domain.

1.7. Current Products

Many products currently on the market offer either no or very rudimentary mobility support. [7] mentions no mobility support for none of the following products: AirLAN by Solectek Co. San Diego; ARLAN by Aironet Wireless Communications Inc. Akron, OH; FreePort by Windata, Littleton MA; InfraLAN by InfraLAN, Acton MA; RangeLAN by Proxim, Mountain View, CA, Maynard, MA. Early versions of the Altair Plus II by Motorola Inc. required the user, once he moved into a new cell, the new frequency had to be programmed into the mobile unit by connecting a terminal to the serial port and reprogramming its frequency. The current version Altair Plus II the mobile terminal can contain a list of authorized frequencies. With *search mode* enabled, the mobile terminal searches through the frequencies specified in its internal list until it finds and locks onto the frequency used by a nearby base station. However no statement is made on the duration of this interruption, the velocity of the mobile terminal that may be supported and on the procedure that is carried out on the interconnecting link between the two base stations.

The RoamAbout access point by DIGITAL is a wireless-to-wired bridge that connects local wireless personal computer networks to a wired Ethernet LAN. The RoamAbout access point enables users to move from the coverage area of one access point into the coverage area of another, providing a clean, seamless handover between two cells that physically overlap each other. To support in-building roaming, each access point learns the nodes of all the wireless hosts within its coverage area that are to communicate with the wired network. Unfortunately no statement is made how fast this can be achieved, what price is paid to keep the information constantly updated and what requirements have to be fulfilled.

In a WaveLAN by AT&T the access points are interconnected via a wired backbone to provide inter-cell communication. Roaming stations that have been moved out of range of a access point will automatically start searching for a new access point that will provide a higher level of communications quality. If needed, the station will automatically sign on to a new access point to maintain its wireless connection. The scheme relies on signal strength measurement of beacons continuously transmitted by all base stations using a good share of the available wireless bandwidth. In order to compare the different signal strength levels both base stations have to operate on the same operation frequency which not only limits the frequency efficiency severely but also is likely to result in frequent hidden terminal scenarios.

1.8. State of the Art / Contribution of this Thesis

There was some work going on in the field of mobile communications at the time of the beginning of this thesis, that were - even though starting from different points and heading into different directions - overlapping partially with the problems described in this thesis. In the Internet environment, mobility management had been developed and implemented for the network layer in the MobileIP protocol[29], which however is found to be of only limited usability in the environment here, as discussed later in Section 2.3. Work on the HIPERLAN-1[4] and the IEEE 802.11[3] standards for physical layer and access control layer in wireless indoor LANs was under way giving way to the explosive growth of WLANs. For the former of the two the concept of forwarding had been introduced but questions remain about its immediate applicability (see Section 2.1.). The term of ad-hoc networking had just been reinvented in the new context of non-hierarchical wireless high-speed data networking however many of its problems were just recognized in their need to be solved first. With the arising of wireless ATM on the horizon, many work went into the question of handover of connections in those networks. [58] gives an overview and discussion about possible handover schemes.

The architecture chosen by the IEEE 802.11 group for their standard for a wireless LAN, as shown in Figure 3 would allow mobile communication within the locally administered range of an extended WLAN. Early in 1997 the removal of an essential part of this concept - of the distribution system (DS) - had been proposed in the 802.11 working group[8], with the argument that the working group was lacking any expertise in defining an interoperable DS at this time. It was considered immature to include such a concept in the standard, and it was proposed to define it as “work-in-progress” and “under study”. The decision taken on this issue was that such “concepts were purposely not defined by P802.11 because the functionality (above MAC layer) would be beyond of the IEEE 802-working group mandate.

As a result of this decision, interoperability between access points of different vendors was (and still is) not possible. Only in mid-2000, more than four years after the start of the work for this thesis, the IEEE 802.11 working group has recognized this problem and has issued an invitation [6] to submit proposals on the subject of a protocol for inter-access point interoperability.

This thesis fill this gap and to discuss the space for possible solutions. It will develop mechanisms applicable for a distribution system, which have not been developed so far. It will contain a systematic investigation of mobility management schemes designed for local area connectionless traffic. As a major part of the original work of the thesis, handover enabling mechanisms will be specified for the IEEE 802.11 environment. In this thesis a number of management schemes within this system have been developed and will be presented together with specifications for them. The schemes will be evaluated with the help of simulative examination. The simulation results will be thoroughly discussed.

The reader of this thesis will find out about the range of possible solutions that can be chosen to solve the presented problem. After reading this thesis the reader may easily decide which of the presented mechanisms will best suit his requirements in his particular environment. Development engineers working on similar problems may get quantitative orientation on the impact of a handover interruption on the packet flows performance parameter.

Chapter 2

Alternate Architectures

2. Alternate Architectures

A number of strategies may be applied to satisfy the functional requirements of a range extension concept as elaborated in the introduction:

- extend LAN range beyond the single radio cell dimension,

- allow / tolerate mobility between the multiple cell areas within the extended LAN

Depending on the scheme the resulting system will offer different features in terms of mobility management and range extension.

The different solutions may be categorized where the additional functionality is located with respect to the OSI 7-layer model. On the MAC-Sublayer of the Link Layer one may apply forwarding and bridging concepts, on the network layer one may achieve mobility support with the help of mobile network protocols like MobileIP ([29],[30],[31],[32]), few solutions work with transport layer protocol extensions or user space solutions ([42],[43],[44]). In the following subsections the characteristics and limitations of each of the solutions will be discussed.

Most of the solutions presented thereafter apply devices that have been developed and are commonly used for interconnection of subnetwork entities (bridges, router, switches). In those cases one will mainly have to evaluate their suitability for support for mobile hosts.

2.1. MAC Layer Approach - Forwarding

Forwarding is a solution designed to extend the range of a WLAN beyond the senders radio range with the help of stations passing packets on across the wireless medium towards the destination host. Figure 5 illustrates a single-hop forwarding scenario.

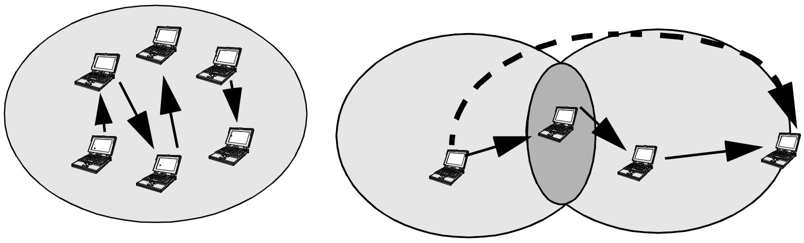


Figure 5:Forwarding: Direct Communication and Forwarded Communication

The Hiperlan standard [4] has integrated this concept, as well as the DECT group of standards that defines DECT relais[10], to extend the coverage area. Figure 6, (taken from [11])shows a scenario with either fixed or mobile DECT relais stations.

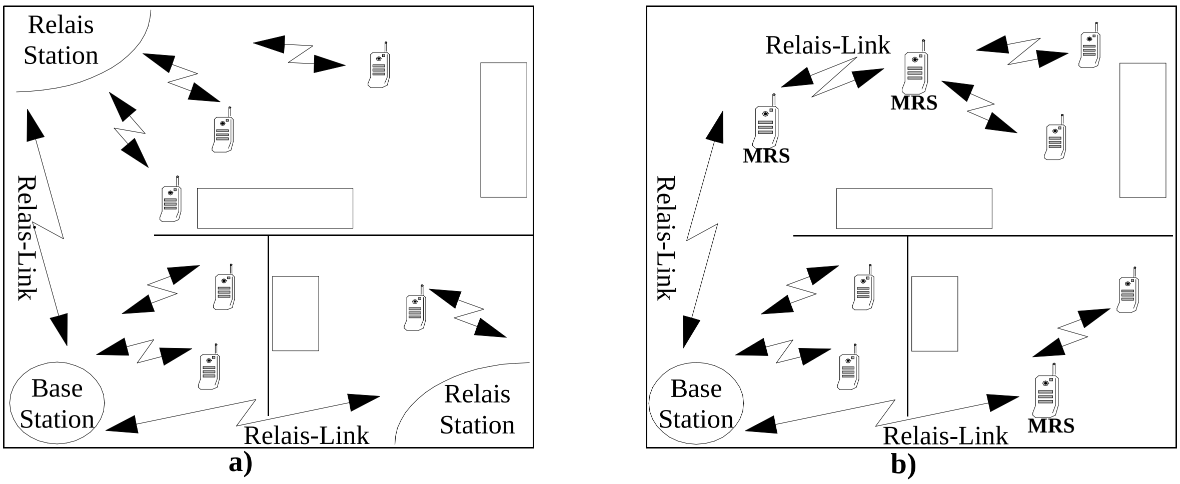


Figure 6:DECT Relais with Fixed Relais Stations (a) and Mobile Relais Stations MRS (b)

Since no fixed infrastructure is *needed*, forwarding is especially well suited for an ad-hoc networking environment. Not every station in the radio cell needs to operate as a forwarder, however each station offering its capabilities as a forwarder needs to know the network topology to a certain degree in order to make a justified routing decision. E.g. in Figure 6b, the Mobile Relais Stations (MRS) need to know, that they have to act as the relais system for some mobile stations, which may not be possible for a longer period of time if the MRS themselves move around, the mobile stations need to find the closest relais station or base stations, and the MRS needs to decide which next station - another MRS or base station - to choose. Further problems arise with the connection-oriented approach as applied in DECT - the chosen path should be valid for the lifetime of the connection or has to be re-setup - or with the connection-less approach applied in Hiperlan. The topology information may be retrieved and maintained e.g. by continuously transmitting and receiving special control PDUs and ageing. If a non-forwarder wants to transmit a packet to a node outside of its radio range it either addresses the next forwarder or broadcasts it to all neighboring stations. Every packet is relayed from forwarder to forwarder until it reaches its final destination either by unicast relaying or broadcast relaying or until its lifetime is expired.

The approach forwarding only offers a solution to communicate with stations outside of the own radio range, i.e. it does not explicitly address mobility related issues. Mobility management can be provided depending on the features of the applied mobile mesh network routing protocol, which is an open research field. The advantage of forwarding however is the instantaneous possibility of using the mechanism without any additional installation of infrastructure. In cases of very low load the implicit hidden terminal scenarios will not be a big obstacle for successful communication beyond radio range.

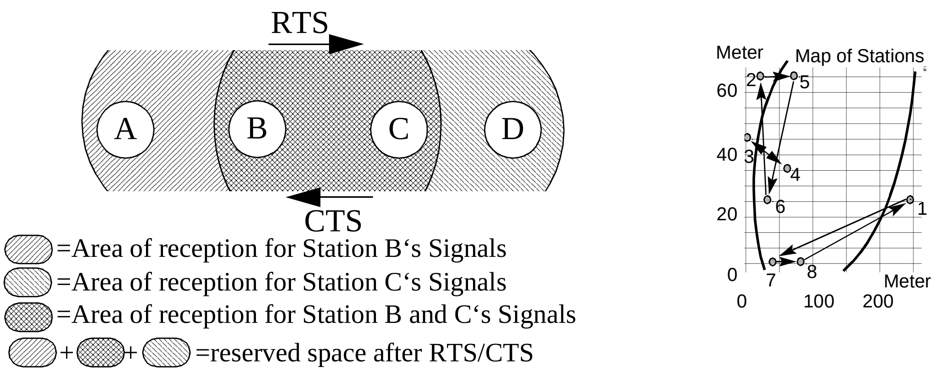
Forwarding introduces a number of new problems. First of all control information have to be exchanged between the mobile nodes in order to update the topology periodically - for an effective routing decision, the forwarder has to have a consistent image of the topology at the very moment. Since common routing algorithms are not designed for the continuously changing network topology new algorithms have to be designed. Within the IETF a group has formed to start working on “Mobile Ad Hoc Networking (MANet)” Currently several groups are researching solutions for this problem ([12], [13], [14], [15], [16], [17], [18]).

Next, the forwarding concept relies on the presence of forwarding stations - mobile hosts that are willing to donate energy and processing power to the benefit of other stations. Therefore in an environment, where these resources are scarce it might be hard to find such a ´volunteer`.

Third, some packets have to travel across multiple *wireless* hops to their destination. As wireless links are known to be more error-prone than wired links, this increased chance of transmission errors increases the risk of errors (and subsequent need for retransmissions or other error correction mechanisms) on the overall path. Also the scarce wireless bandwidth will have to be used multiple times for a single transmission. From the internetworking point of view in order to support mobility one has the major advantage that no re-routing of the end-to-end-connection is needed. The relaying of data units to the destination is done invisibly within the wireless LAN. However since the forwarders may be mobile as well the forwarding chain may break at any time leaving the destination unreachable. Problems like continuous service and hitless switching are inherent features of this approach as long as the dynamic forwarding algorithm works appropriately.

2.1.1.Hidden Terminal Problem

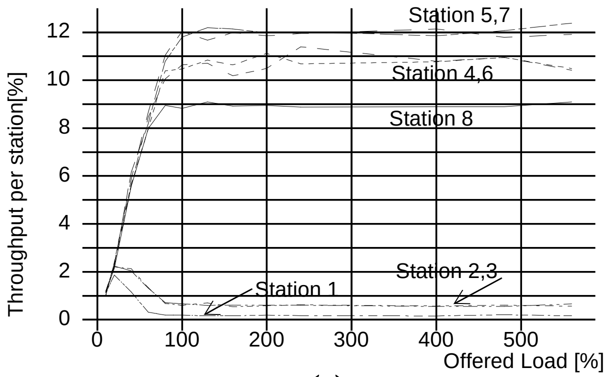
However the most serious problem with the application of forwarding stems from the hidden terminal scenario - a problem dependent on the wireless medium, that requires special attention on the MAC design in a wireless environment. The scenario is briefly described here with reference to Figure 7a:



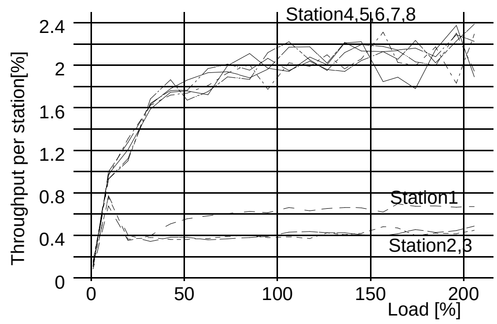
(a) **(b)**

Figure 7:Hidden Terminal Scenario (a) and simulated station topology (b)

If station B is sending to station C the medium appears busy only for the stations located within the range of the sending station B - other stations cannot sense a signal and consider the medium idle. Therefore, station D might start a transmission since it does not notice B‘s ongoing transmission. However, since the receiving station C is within range of B *and* D and thus receives two signals at once it will not receive any undisturbed signal (whether destined for it or not). However, this collision cannot be detected at the sending station B unless it notices the lacking acknowledgment from station C after a certain time-out. Both protocols selected for access in WLANs - IEEE 802.11[3] and Hiperlan[4] - are equally vulnerable to this scenario, as shown in the following simulation results. The simulation setup (Figure 7b) consists of a cluster of 5 stations, that each are within range to receive the signals of all 8 stations present, and one station (number 1) being hidden to the remaining two stations (2 and 3). The curves in Figure 8a and Figure 8b show the decrease of throughput for each of the stations with increasing offered load for IEEE 802.11 and Hiperlan respectively, if hidden terminals are present. These reslts are taken from previous work. An extended and more detailled description of the simulation, the parameters and further results can be found in the previously published publication [22].



(a)



(b)

Figure 8:Throughput Hidden Terminal Setup IEEE 802.11(a) and Hiperlan (b)

In both WLANs one can observe different behavior depending on the position of the sender and the receiver. The simulations for 802.11 (Figure 8a) show that stations that send towards hidden stations (4, 6, 8) and the hidden stations themselves (1 and 2,3) achieve significantly lower throughput than the other stations (5,7). The first group (4, 6, 8) gets packets through successfully, however many acknowledgment packets are destroyed by traffic from the hidden stations. The breakdown of inbound data traffic in the case of higher load stems from the fact that the mutually hidden stations become synchronized by an earlier data exchange in the area between them. As a result, they start their backoff counters at the same time but they are unable to detect the begin of transmission of the other station.

A similar negative influence occurs in Hiperlan-networks (Figure 8b). The simulation shows the decrease of the overall throughput in that scenario. With increasing load the damage done by the hidden terminals to the overall load increases. Even the peak achievable throughput at 40\% load is significantly lower than the throughput without hidden terminals. The Hiperlan draft standard does not yet attack this problem.

2.1.2.Hidden Terminal and RTS/CTS in IEEE 802.11

The IEEE 802.11 group realized the necessity to address this problem and integrated the RTS/CTS mechanism, developed in [25] and analyzed in [26] to solve it: Each station competes for access as described in the general access procedure. When the RTS/CTS mechanism is applied, the winning station does not send data packets right away but sends a RTS (Ready To Send) packet to the receiving station, that responds with a CTS (Clear To Send) packet (see Figure 7a). If a station captures a RTS packet from another station and it is not the destination of the RTS packet it reads the intended transmission duration from the RTS packet and stays silent for that time. The same happens if only a CTS packet is received i.e. by a station outside of the transmission range of the sender but within the range of the receiver. This guarantees that all stations within range of either sender or receiver have knowledge of the transmission as well as of its duration.

The effects of the RTS/CTS mechanism are as follows

- It increases bandwidth efficiency by its reduced collision probability since the ongoing transmission has been made known everywhere within the range of it
- It increases bandwidth efficiency since, if collisions occur, they do not occur with the long data packets but with the relative small control packets
- It decreases bandwidth efficiency since it transmits two additional packets without any payload
- It decreases bandwidth efficiency since it reserves geographical space for its transmission where or when it might actually not need it.

The same setup as in Figure 7b has been used for the following simulations, however **with** the RTS/CTS mechanism. As can be seen in Figure 9, the RTS/CTS message exchange does not completely solve the hidden terminal problem, even though significant improvements can be achieved. Still the stations that are hidden to other stations hardly get any packets through due to the above mentioned synchronization effect. Station 1 still hardly gets any packets through, but its throughput is improved compared to the figure without RTS/CTS. The same goes for stations 2 and 3 - all of the hidden stations benefit from the captured CTS packets. The non-hidden terminals all achieve the same (high) throughput due to the fact that outbound traffic is protected by the RTS packets. This simulation has been published and thoroughly discussed in the previous publications [20], [21] and [22].

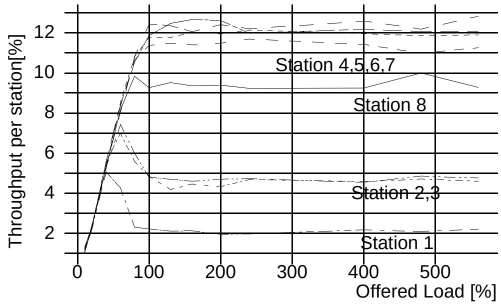


Figure 9:Throughput Hidden Terminal Scenario RTS/CTS in IEEE 802.11

2.1.3.Hidden Terminal and Forwarding in Hiperlan

In order to investigate the influence of the hidden terminal scenario on forwarding as defined in the Hiperlan standard, a scenario as depicted in Figure 10 has been investigated.

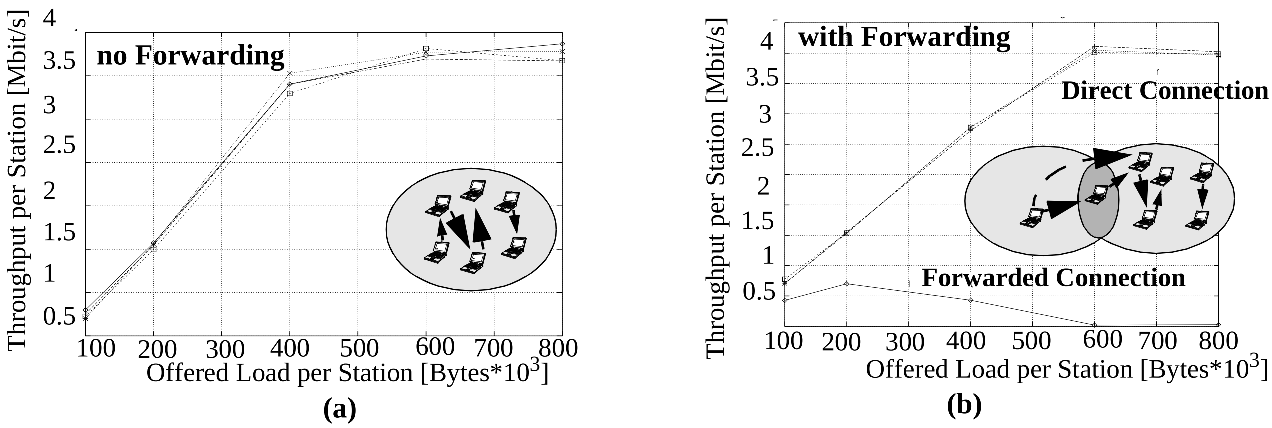


Figure 10:Performance without (a) and with (b) Forwarding in Hiperlan

The simulation examines traffic from a single terminal which is outside radio range of a second cluster. Please note that such artificially terminal configuration has been chosen in order to stress the problems one has to be aware of in the case of forwarding. The simulation results show that an isolated station almost has no chance to send into a cluster of communicating stations due to the hidden terminal effect. In fact starting with fairly low overall load the throughput from the hidden station into the cluster decreases rapidly with increasing overall load. The reverse case - traffic to a single terminal outside a cluster - causes similar unsatisfying results. It can be observed that the isolated station is able to receive packets from the cluster even under high overall offered load. But one can also see that the achieved throughput per station is lower than in the case of a fully meshed network with the same overall offered load. A more detailed discussion of the simulation results can be found in [24].

Looking at the approach 'forwarding` , one can say that it may be able to fulfill the task of extending connectivity beyond the radio range, however it relies on the presence of forwarding stations - mobile hosts that are willing to donate energy and processing power to the benefit of other stations. Strategies to selecting stations for this resource intensive special task may be developed, e.g. stations having access to a power outlet may be primary candidates here. In cases of very low load the implicit hidden terminal scenario will not be a big obstacle for successful communication beyond radio range (an example scenario for this case is an in-house ad-hoc network between electronic appliances, where the occasional messages to e.g. the toaster are forwarded from the laundry machine). With increasing load however communication for a host outside of a cluster becomes almost impossible. Lately a new access scheme has been published that promises to solve the hidden terminal problem to some degree [27]. It remains yet to be shown whether this is true in an environment as applied here under all circumstances (e.g. a hidden terminal scenario may not have been present at the time, when the mechanisms to manage the hidden terminal scenario are taking place, however a mobile host may move into the area and become a hidden terminal at any time after those mechanisms are completed). Once ad-hoc routing algorithms have been developed, once the hidden terminal problem will be solved, once better batteries have been developed, forwarding may become a useful approach. The advantage of forwarding is the instantaneous possibility of using the mechanism without any additional installation of infrastructure, which also results in a lower cost per interface. If network addresses are considered a scarce resource, as it is the fact in the present IP version 4 protocol, this approach is undesirable to waste IP subnetwork addresses since each cell remains a full single IP subnet.

2.2. MAC Layer Approach - Bridging

Bridges[19] were originally designed to interconnect a number of LAN-segments and as such seem to be a natural solution to interconnect wireless cells (=subnets) as well. Bridges are a readily available technology operating on the data-link layer. Bridges analyze incoming frames, make forwarding decisions based on their forwarding tables and information contained in those frames, and forward the frames to their destinations. A scenario using bridges interconnecting wireless subnets across a distribution subnet is shown in Figure 11. E.g. [28] proposes such a bridging concept for interconnection of wireless LAN cells with a self learning routing table and an ageing scheme where routing entries are deleted after a certain number of seconds.

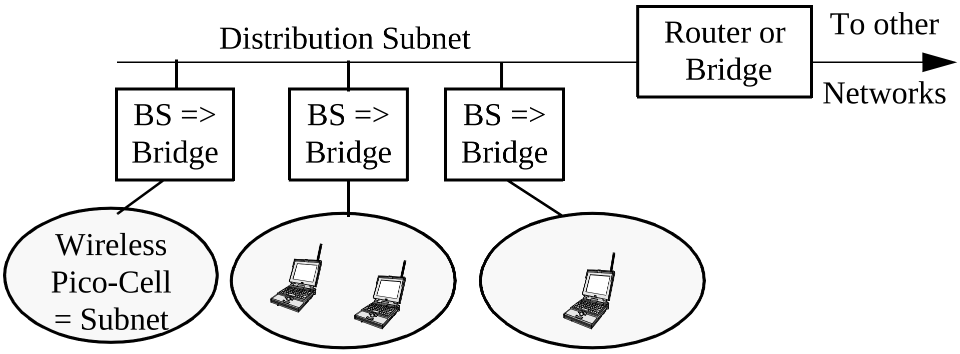


Figure 11:Using Bridges to Interconnect Wireless Cells

The critical point in this approach however is the lack of support for mobility. In order to operate correctly, each bridge (i.e. BS) needs to know the current location of the MHs as well as the corresponding BS. In the static configuration, that bridges are designed for, they get this information by apply different algorithms, e.g. the spanning tree algorithm for transparent bridges or discovery packets for source routing bridges. The spanning tree calculation occurs when the bridge is powered up and whenever a topology change is detected. Configuration messages called Bridge Protocol Data Units (BPDUs) trigger the calculation. These messages are exchanged at regular intervals. Bridges exchange configuration messages at regular intervals. If a bridge fails (causing a topology change), neighboring bridges will soon detect the lack of configuration messages and initiate a spanning tree recalculation. This however does not solve the case of mobile hosts - the base stations remain fully operational if hosts move around. Another mechanism based on ageing is needed to take care of configuration changes on the different subnets. Entries are periodically removed from the hash table in bridges if they have not been refreshed for a few minutes. In this way, if a computer is unplugged from its LAN, moved around the building, and replugged in somewhere else, within a few minutes it will be back in normal operation, without manual intervention.

As a rough estimation, assuming an update frequency of one update every 4 minutes, as it is common to bridges, and a linear motion of the mobile host with the speed of 1 meter per second for the mobile host, between 60% to 95% of the time (for cell diameters of 100 and 10 meters respectively) the mobile host will already have left the cell, that has been listed in the update as its current location, i.e. outdated information is kept in the bridge tables. Increasing the update frequency will cause more configuration message traffic and will require more processing power from the base stations. Therefore one can say that todays bridges cannot be applied of-the-shelf to solve the interconnection and location issues arising with the chosen architecture from Figure 3, they are not designed for management of mobile stations that move around fast between wireless subnets. If only quasi static service is required this scheme serves well at the expense of mostly giving up one of the main additional features of wireless networks - i.e. mobility. The applicability of a bridge-like algorithm however - a periodic table exchange to distribute location information - will be investigated later on.

2.3. Network Layer Approach - MobileIP

The MobileIP working group of IETF has developed a series of Internet Drafts and Internet Standards (Request for Comment - RFC) to define a protocol allowing mobile nodes, usually portable computers, to move around while maintaining internet connectivity. These activities form the first approach to standardize mobility management in an internet environment. The original drafts gradually evolved into a few branches, namely, IP Mobility Support[29], IP Encapsulation within IP [30], Minimal Encapsulation within IP [31], and Route Optimization in Mobile IP [32].

The chosen approach uses a home agent in the home network that catches and tunnels incoming packets towards the foreign agent in the visited network. The foreign agent then forwards the packets to the mobile host. An scenario for MobileIP is shown in Figure 12.

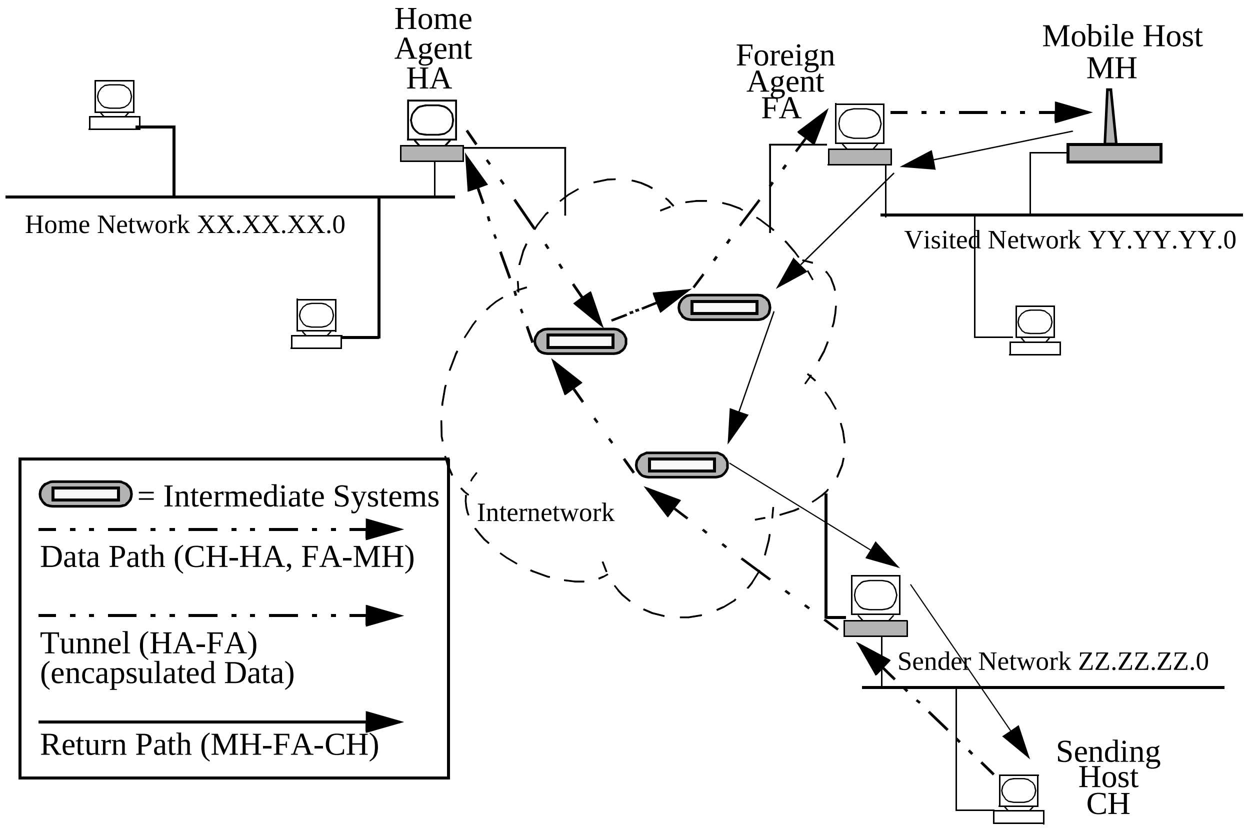


Figure 12:General MobileIP Scenario

With its ability to accommodate mobile hosts in an internetwork one might also consider applying MobileIP in the E-WLAN scenario of interconnected wireless cells. Such a setup would require that each cell forms a full IP network, i.e.one would not interconnect MAC-layer subnets but IP networks with an own IP network address. This requires assignment of multiple IP addresses as well as administration of multiple networks instead of just one. Each mobile host would have to be assigned a "*home agent*" in its "*home network*". If the mobile host visits another cell, any data destined for it would be routed to the home network where the home agent would intercept the packet and tunnel it to the foreign agent at the mobile's current location - just as in the above described general scenario. Figure 13 shows such a setup of multiple cells interconnected with a MobileIP router. Each shaded area represents a fully autonomous IP network.

As one can see in the figure below, each packet sent from one cell other than the home cell of the receiver to this mobile receiver that is away from home, needs to go across two or three (if a direct connection between the two base stations exists) additional hops. These additional transmissions even in the static case have direct impact on the delay characteristics. [33] describes measurements (in a similar setup with mobile hosts however without wireless links) that show that the use of MobileIP in their environment causes ~3.5 times the transmission delay compared to an environment without MobileIP and mobility..

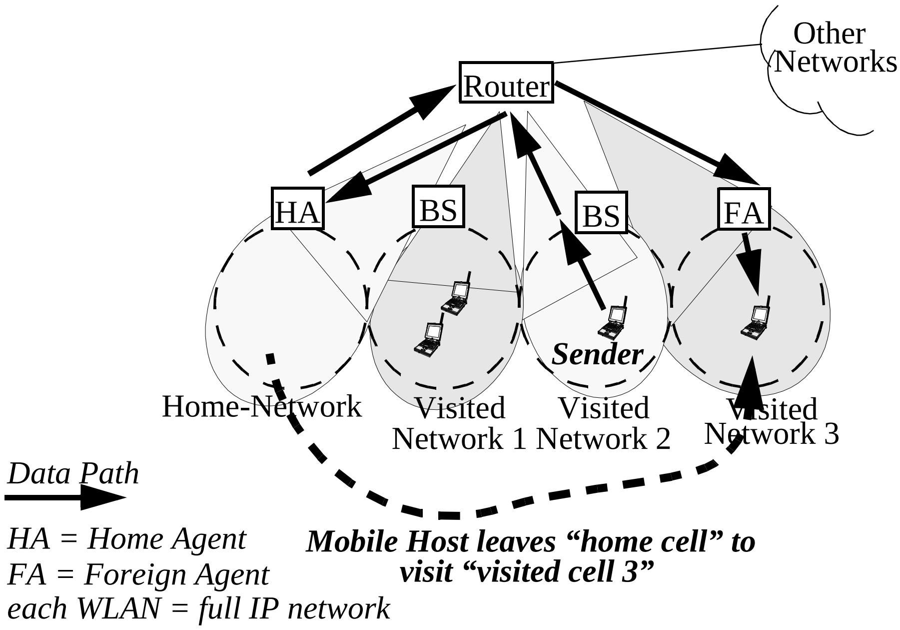


Figure 13:Using MobileIP to Interconnect Wireless Cells

If the mobile host is visiting from a remote network the delay would increase even more since its home agent would remain in this remote network and all traffic would have to go via the home agent across eventually multiple long distance hops. Even if this “*dog-leg-routing*” would be avoided as in[32], still the registration and authorization messages as well as the initial messages of the transmission would have to be transmitted to the home agent thus making the reachability of the home agent a single-point-of-failure.

After a mobile host moved into a new cell it will have to solicit a router advertisement, reply on it with a registration request and wait for the registration reply. If the home agent is located far away this time will increase. Additionally some time might have to be spent on configuration parameter exchange either manually or with the help of autoconfiguration protocols ([46], [47]). If all these delays are added together one can say that such interruptions, that are already occurring in the local network are not acceptable for continuous or even semi-continuous service in a mobile multimedia environment.

In acknowledgment of this possibly long delay between cell handovers interconnected with the help of MobileIP, two projects developed improved schemes - [34] and the Daedalus project ([35], [36]). Both schemes introduce a second layer for local area handover management in order to reduce the number of '*expensive*' MobileIP handovers. The first scheme uses local rerouting from the old base station to the new one. The latter scheme applies a local area multicast tree to distribute data for the mobile host to all base stations in the neighborhood. Each handover may now be managed locally since the multicast area spans over multiple cells and the mobile host can receive its data at each base station within this area.

These schemes however have a number of disadvantages and problematic approaches. First the approaches share the same drawback as plain MobileIP with its integration of multiple full-blown IP networks. Second both schemes require that the mobile host listens to periodically transmitted beacons from the base stations, possibly even multiple beacons from different base stations at the same time and that it is able to determine the closest base station with the help of signal strength measurement facilities. These requirements however are not easily provided by network interfaces. Once the interface card is either sending or transmitting it would need a second antenna in order to continue to receive the beacons from the base station. However neither this second antenna nor the signal strength measurement facility is available on most of the currently available interfaces. If they would be added they would cause significantly higher cost-per-interface. In order to be able to evaluate two signals at a time the cells necessarily need to overlap for a certain area and must not have idle space between them. In WaveLAN, where the signal strength measurement for beacons are available, the application of it requires that both (or more) cells operate on the same frequency in order to be able to receive the beacons from both (or more) base stations leaving both (or all) the cells with the bandwidth of just one cell. With the undesirable frequent transmission of beacons, that are further reducing the already scarce bandwidth of the wireless link, one can say that these approaches does not offer a feasible solution for a **local** area mobility management with **frequent** handovers.

2.4. Switching / VLAN Concept

Lately LAN switching has become a popular technology to connect several segments of a LAN offering each segment the full medium bandwidth. They apply a similar if not identical concept as in bridges and therefore offer similar features for the chosen scenario. Figure 14 shows an example setup of wireless cells interconnected with a LAN switch.

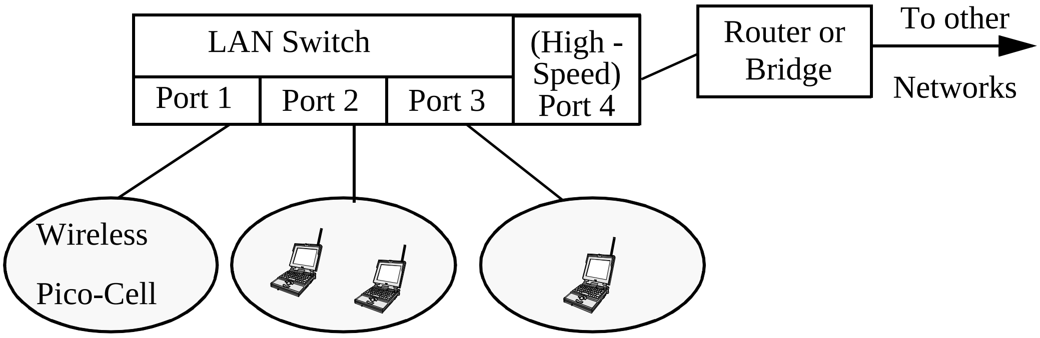


Figure 14:Using LAN Switching to Interconnect Wireless Cells

A solution applying LAN switches to interconnect wireless cells would have the advantage that hardware is readily available and that a port to other networks can be adapted to the bandwidth in those networks. As with the before discussed approach based on bridges the approach offers interconnection service however is not specifically designed to support mobility. The purpose of LAN switches is to provide more bandwidth for each host attached to the medium, but they are targeted primary for a fixed configuration with only rare reconfiguration events. Therefore most LAN switches require extensive manual configuration, almost no dynamic or automatic configuration algorithms are implemented. Each mobility related reconfiguration would therefore cause interruptions of uncertain length. In addition to this LAN switches and the accompanying 'Virtual LAN - VLAN'-concept suffer from the fact that they are only proprietary solutions - no standardization of such devices has yet taken place. Each product may therefore show different features and behavior.

2.5. Mobility Management in GSM and other cellular networks

Similar to the MobileIP approach the widely used GSM technology for cellular telephone networks applies two registers containing location information: a Home Location Register (HLR) that is used to contain the user's location and subscribed services and a Visitor Location Register (VLR) is used to track the location of a user. As the users roam out of the area covered by the HLR, the mobile station notifies a new VLR of its whereabouts. The VLR in turn uses a separate signalling network to signal the HLR of the mobile station's new location. Through this information, mobile terminated calls can be routed to the user by the location information contained in the user's HLR.

GSM networks have a number of different properties that make this technology unsuitable for the chosen architecture: first it provides connection oriented service; second it provides considerably lower data rates (<64Kbit/s) that are sufficient for voice transmission; third it operates on rather large cells of several kilometer and therefore does not need to manage frequent cell handover; fourth it relies on the existence of a certain overlapping area between the cells that allows signal-strength-measurement initiated handover; fifth, its approach based on central databases scales only to a limited number of hosts and database requests, i.e. it limits the mobility parameters that may be supported.

Newly added data services in GSM networks like GPRS for GSM networks or - for the US AMPS counterpart - EDGE or CDPD build upon the handover mechanisms found in the connection-oriented parts of the wireless networks [5]. In GPRS a MobileIP-like handover with home agent (here GGSN - gateway GPRS support node) and visited agent (here SGSN - serving GPRS support node) is applied. The mobile requests a routing update, once it detects a weakening signal. The network then sets-up a new tunnel from the GGSN to the new SGSN. The handover decision can be taken either in the base station or in the mobile station itself.

In CDPD, similarly the handoff mechanism is based on the MobileIP principle. The reason for this is that the HLR/VLR concept, found in cellular networks, maps nicely into the MobileIP concept of HA and VA. In CDPD the mobile station monitors the quality of beacons transmitted from the base station and decides on the handover. The method can thus be classified as MCHO - mobile controlled handoff. After making connection to a new base station the tunnel from the "mobile home function MHF" is redirected to the new base station.

Both GPRS and CDPD operate with tunnels, that provide a virtual connection from the home agent towards the currently serving base station. This quasi-connection-oriented approach is advised since a large (connection-oriented underlying) network cloud is operated thus both end points are under full control of the network operator. In a fully connection-less environment different mechanisms need to be applied. Also the comparably low data rates permit longer interruptions without a too high degree of data loss in the intermediate period. Therefore the disadvantages, that have been elaborated in Section 2.3. do not fall into account here.

2.6. Mobile ATM

Similar properties are likely to be found as well in future wireless ATM networks[40]. The connection oriented end-to-end semantics demands connection rerouting and explicit location registration, update and tracing that involves central location tables in the ATM network. Wireless ATM (WATM) islands should be able to offer seamless handovers with ambulant speeds within the customer premises network (one domain). There are already many ways to implement wide area mobility management, one of the existing techniques could be adopted. The GSM solution explained above solves the problem, another way to implement roaming and wide area mobility service is for example to base them on the advanced intelligent network technologies. Proposals for handover management in WATM describe virtual circuit extension (with or without loop removal), total rebuild of the VC, partial rebuild to a fixed anchor switch or to a dynamically selected cross-over switch, or multicast to neighboring base stations.

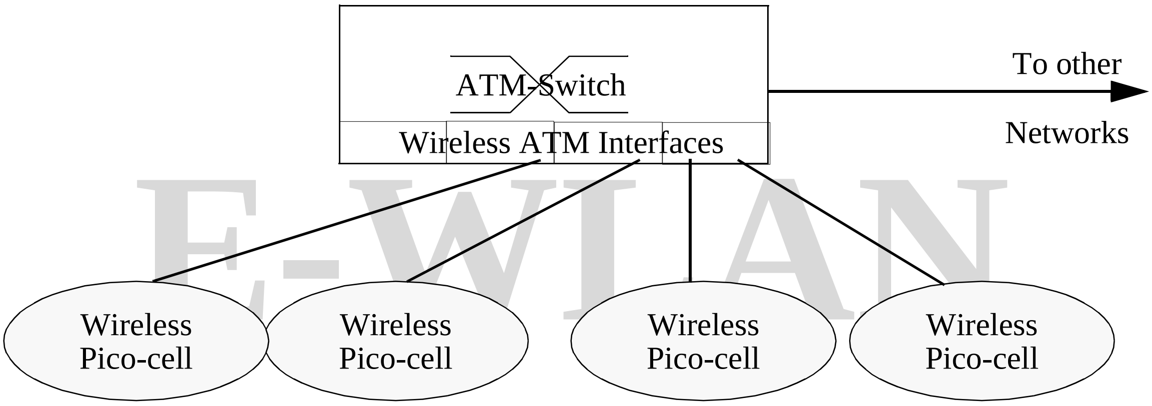


Figure 15:Using ATM switches to Interconnect Wireless Cells

Regarding applicability of such solutions for the architecture discussed in this thesis one has to say that the differences in the environment with its relatively high fixed cost of virtual circuit setup and its connection orientation makes the solutions appear not flexible enough for application in an internet-like environment with variable length messages in a connection-less architecture.

2.7. Transport Layer Approach

The functionality for mobility management may as well be implemented on the transport layer. A group at Rutgers University developed an indirect approach for handling of mobility and unreliability of wireless links on a multihop TCP connection. The approach described in [41] and [42] - even though developed primarily to improve TCP performance over wireless links - also specifies a handover procedure to keep the TCP connection alive once the mobile host switches between cells during the lifetime of the connection. This I-TCP approach uses a separate transport protocol for the wireless link to run a separate transport layer connection between a mobile host and its base station. This prevents the different communication characteristics of the wireless link on the last (or first) hop from having disturbing influence on the remaining part of the multiple-hop end-to-end connection. Since the I-TCP approach requires transport layer processing at the base stations it is possible to accommodate mobility management on this layer as well.

As shown in Figure 16, in I-TCP, after the mobile host has reregistered at the new base station (Steps 1, 2, 3), the transport layer connection is rebuilt at the new base station, then the old base station is informed (Step 6) and sets a forwarding pointer, that allows that TCP packets are redirected on the fly. All I-TCP sockets active at the old base station on behalf of the mobile host must be moved to the new base station (Steps 9 and 10). In addition to transferring the state maintained by the socket and TCP layers there, it involves restarting the connection at the new base station. Also, this migration of sockets needs to be completely transparent to the fixed host at the other end of the connection. The TCP segments that are in transit during a handover are buffered at the new base station even though they cannot be immediately processed. This is necessary to avoid congestion on either side (fixed or wireless) of the I-TCP connection. After an I-TCP handover, the TCP retransmission timer on the wireless side is reset so that the MSR immediately begins a slow start.

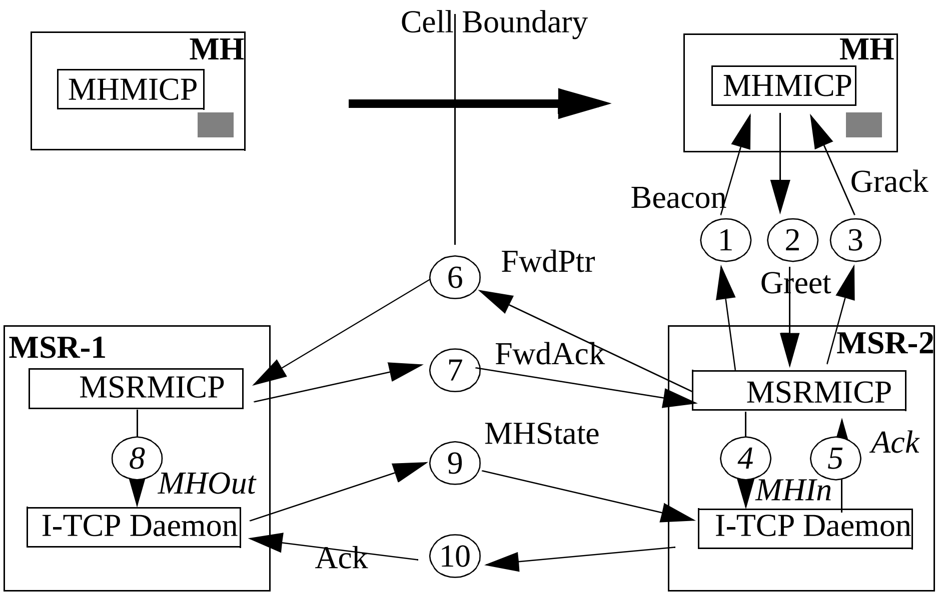


Figure 16:I-TCP Handoff Sequence

Differences occur in the scheme if the cells overlap or if they do not overlap. In the latter case if no direct communication is possible between the mobile host and its previous base station after switching cells the interruption lasts from the moment the old cell is left. In case of overlapping cells, the mobile host can continue to receive IP packets during the time spent in the overlapping area while it sends the outgoing IP packets through the new base station. The I-TCP handover thus does not interfere with other IP traffic to and from the MH. Once the sockets and state information has to be moved the I-TCP connections experiences a brief interruption in the traffic. The TCP segments in transit during this short period are buffered (without processing) at the new base station and are acknowledged as soon as complete state information is available for I-TCP connections at the new base station. With non-overlapping cells, the MH can start sending out IP packets immediately after reregistration, but it cannot receive any IP packets until the rest of the network knows about its new location. This disruption in the network layer is inevitable with non-overlapped cells. [43] writes about I-TCP, that this approach, like other split-connection proposals, in its attempt to separate loss recovery over the wireless link from that across the wireline network, forces every packet to incur the overhead of going through TCP protocol processing twice at the base station (as compared to zero times for a non-split-connection approach), although extra copies are avoided by an efficient kernel implementation. Another disadvantage of this approach is that the end-to end semantics of TCP acknowledgments is violated, since acknowledgments to packets can now reach the source even before the packets actually reach the mobile host. Also, since this protocol maintains a significant amount of state at the base station per TCP connection, handover procedures are likely to be complicated and slow.

2.8. Inter Access Point Protocol IAPP

Aironet Wireless Communications, Lucent Technologies, and Digital Ocean have agreed to develop an IEEE 802.11 compliant interoperability protocol, the Inter Access Point Protocol, IAPP [44] that defines how access points from different vendors communicate with each other to support mobile stations roaming across cells. The IAPP specification builds on the capabilities of the IEEE 802.11 standard. Where the IEEE 802.11 standard addresses the physical and media access control layers of the OSI reference model, the IAPP specification tackles higher-level OSI layers such as logical link control that facilitates inter-access point communications. The protocol is designed to define methods which facilitate interoperability among IEEE 802.11 compliant devices. The IAPP is implemented on top of IP and is an extension to existing management protocols. The protocol stack for IAPP can be seen in Figure 17.

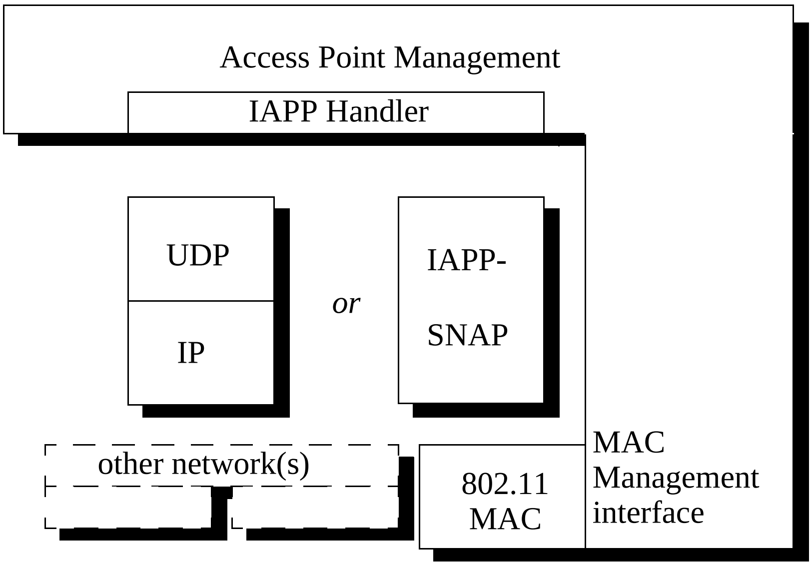


Figure 17:IAPP Protocol Stack

Stations may move from one BSS to another and maintain their higher level network connections. This mobility is enabled through the use of IEEE 802.11 management frames and an Inter Access Point Protocol that is used to communicate on the distribution system (wired or wireless). In order to allow operation over a wide range of networks, two transfer protocols will be implemented: UDP/IP and 802.2 Sub-Network Access Protocol (SNAP). Whenever an IP address is present in the access point, the UDP/IP will be used. SNAP will be used in cases where an IP address is not assigned to the access points desiring to communicate.

IAPP consists of two modules - the announce protocol and the handover protocol. The intention of the announce protocol is to inform the other access points that a new access point has become active, inform other access points and/or network management functions of the continued operation of a particular access point, inform the access points of network-wide configuration information, allow for a network management function to provide AP coordination. This part of the protocol works similar to a router discovery protocol [45] in combination with an autoconfiguration protocol like DHCP [47] or BOOTP [46].

2.8.1.IAPP Handover Procedure

The second module - the handover protocol is used to inform the old access point that a station is taken over by another access point; to release the resources that the old access point had assigned to support the station, update its filter table to forward frames destined for the station appropriately, discard or forward frames buffered in the old access point for the station, update filter tables of intermediate MAC-bridges to forward frames destined for the station appropriately.

The Handover procedure is tied into the IEEE 802.11 Reassociation procedure; it is initiated when an access point receives a Reassociation-request from a mobile station. The Reassociation messages between mobile station and access point (Reassociation-request and Reassociation-response) are part of the IEEE 802.11 MAC protocol. Figure 18 shows the packets that are exchanged after the reassociation between the new and the old base station.

The figure depicts the Reassociation-response being delivered to the mobile station before the HANDOVER.response is received by the new AP. The timing for sending the Reassociation-response is implementation dependent which allows for the Reassociation-response to be delivered after the receipt of the HANDOVER.response. This allows for full implementation flexibility.

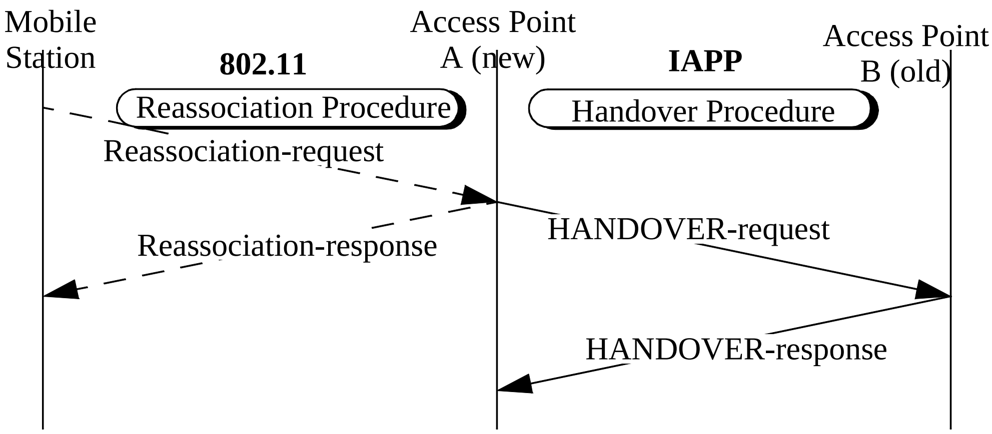


Figure 18:IAPP Handover exchange

From the time the Reassociation-request is sent until receipt of the Reassociation-response, the mobile station must not send any data frames. If during the Handover procedure a (retried)

Reassociation-request is received from the (same) mobile station, the 802.11 Reassociation procedure is completed normally, but no additional HANDOVER sequence will be initiated.

The new access point (A) is also responsible for recovery of the Handover procedure. If the HANDOVER.response is not received within a certain time as established by the Handover timeout, the HANDOVER.request is retransmitted over the distribution system. If no HANDOVER.response is received at all (after x retries), the Reassociation procedure is completed by the new access point. This behavior ensures that if the old access point is not reachable (it died, got disconnected, is on the other side of a router or gateway, or the distribution system is temporarily out of service), the MAC services of the mobile station can continue.

As can be seen from the above description the IAPP protocol is mainly developed to provide facilities for interoperable interaction between base stations from different vendors. This is done by defining few messages that are used to indicate a handover after it has occurred to the old base station, and to initiate table updates in all base stations. No attempt is made to minimize the duration of the interruption of the connection or the number of dropped packets. No clear statement is given whether the packets that were on the way are dropped or whether they should be forwarded. Sending the control information over IP/UDP would again result in multiple full IP networks that are interconnected. By designing a new SNAP protocol especially for IAPP the design essentially follows the architecture chosen for this thesis, however the IAPP protocol only applies its new SNAP version to transmit control messages to the IAPP handler located above the protocols and does incorporate any sophisticated functionality into the IAPP-SNAP. The transmission of data messages over the IAPP system is not specified. However the general framework of the IAPP-SNAP protocol would also allow to integrate much more sophisticated solutions of a distribution system as will be presented in the following chapters.

Chapter 3

Distribution System - Architectural Discussion

3. Distribution System:

The approach chosen in this thesis to achieve interconnection of multiple wireless cells and mobility management applies a wired backbone by adding functionality for a distribution system on data link layer. The goal is to offer unconstrained support for mobility for the mobile hosts among the connected pico-cells. As discussed in the previous chapter a solution between the MAC layer and the network layer offers advantages compared to solutions operating on other protocol layers. Unlike a forwarding approach it does not use additional bandwidth on the wireless link; unlike bridging or switching on the MAC layer it may be capable to allow for fast adaptation to changing configuration; unlike network layer approaches like MobileIP or transport layer approaches it does not use one IP subnet per cell, does not require long distance mobility management and allows local area transparent mobility support.

The architecture as shown in Figure 3 consists of wireless picocells where mobile hosts roam around, one base station in each cell offers access to the wired distribution system. The distribution system itself offers access to shared resources like printer or file server and to the portal that connects the whole E-WLAN to other networks.

Such an architecture glues together separate wireless communication islands into ONE network. It allows for two hierarchical layers of mobility management - most handovers between cells can be managed in the local domain, only if the mobile host leaves the local E-WLAN and moves into another administration domain one needs to perform a network layer based handover. The frequent local area handovers can be handled much faster and more efficient. The second type of handovers will now occur far less frequent and therefore its disadvantages (long delay, long interruption, single-point-of-failure, dogleg routing) do not fall into account. Also since most traffic to or from outside networks comes in via ONE device (typically a gateway, router or firewall). In the case of local area mobility of the receiver it does not seem advised to involve any station beyond that entrance hub in the mobility management since no changes will occur there (Figure 19). This architecture is well suited to build up a building-size wireless infrastructure.

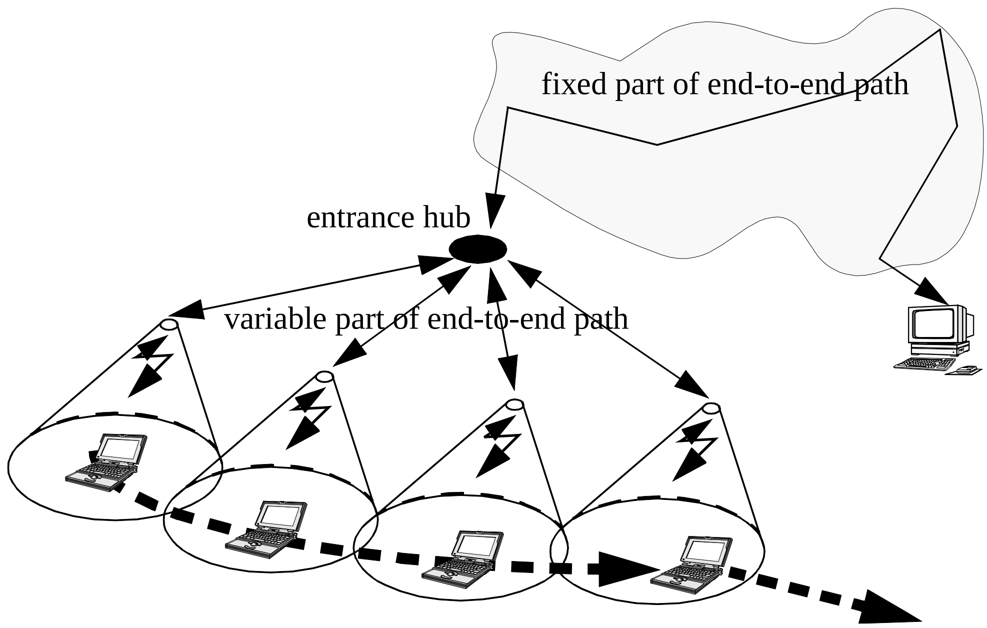


Figure 19:Fixed and Variable Part of End-to-End Path in the Case of Local Area Mobility

3.1. Metatask / Functional Requirement

The primary goal of the chosen architecture is to provide a *LAN-like feeling* for all participating hosts. *LAN-like feeling* means that a mobile host should not experience significantly different communication behavior in terms of delay, throughput, connectivity and available services as if it would be attached to a wired LAN. This should remain valid as far as possible even in the presence of mobility. The resulting E-WLAN should have comparable properties as commonly used LANs in terms of coverage area, number of supported hosts, transmission speed and service quality. Any added functionality must remain transparent for corresponding hosts outside of the E-WLAN since those hosts will have no means to detect the reason for possibly different behavior and are thus likely to react in a counterproductive way. One can assume that the available bandwidth of future wireless communication will remain being smaller compared to wired communication. Therefore as far as possible the solution should not increase the overhead on the wireless link.

3.2. Subtasks

The problem of designing a distribution system with the above described features splits up into two main subtasks - mobility management and packet transport.

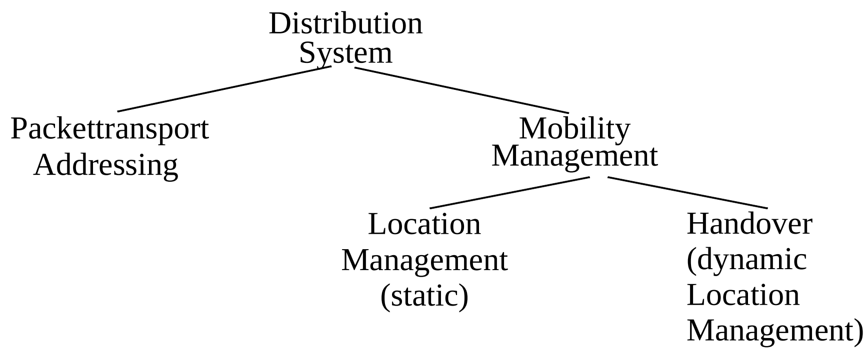


Figure 20:Subtasks of a Distribution System

Mobility management consists again of the subtask to create an up-to-date [host: location]-table and the subtask to manage handovers. Once the location is known means are needed to transport the packet to the destined host, i.e. one needs to enable *location dependent addressing*. In the following sections I will discuss both these subtasks.

The thesis mostly describes solutions for a communication where both stations - sender and receiver - are located within the E-WLAN. The case where the receiver is located outside the E-WLAN does not impose any additional activities in the local E-WLAN, traffic simply is forwarded out to the destination by the portal by using common routing techniques. If the destination outside the local E-WLAN is mobile and changes its location, the mobility management will have to be carried out at that outside network without intervention of the senders network. If the sender is a host outside of the local E-WLAN the portal acts as the “base station” of the senders “cell” (which in fact is the internet) to the inside, and as a gateway or a router to the local (E-W)LAN to the outside. The main problem in that case is that the sender outside of the E-WLAN cannot be expected to participate in the mobility management scheme. Besides that the larger delay on the path between sender and the portal has to be considered.

3.2.1.Underlying Wireless Network

In the following an underlying wireless network will be assumed that is offering WLAN-typical services like association/reassociation/disassociation and features like immediate acknowledgment. The former services are necessary even in a stand-alone WLAN and their implementation is therefore outside of the scope of this thesis However the distribution system will build upon those services on some occasions. The presence of the feature immediate acknowledgment can equally be assumed since it is generally applied in WLAN implementations - the vulnerability to the unstable medium characteristics requires stricter data integrity protection compared to the wired counterparts.

3.2.2.Mobility Management

As stated above location management needs to determine where - as a function of time - is the correct destination for a packet? The goal is the creation of a currently valid [host: location] - table. To start off easy I will first look at a non-mobile configuration.

3.2.2.1. Solving the Static Case

Some procedures have to be carried out by the MH and BS visited at first at the time when the mobile host initially enters the E-WLAN. First either the MH needs to discover that there is a BS offering its services in this area or the BS must detect the MH requesting these services. Next a couple of procedures have to be carried out namely host- and user registration, host- and user authorization, temporary IP-address assignment and other LAN configuration tasks. Except for the additional procedures for host authorization (which is usually unnecessary in a wired network where the necessity to connect physically to the wired medium usually prohibits unauthorized hosts from accessing the network) this is comparable to any other LAN where these configurations are carried out manually on central server and the newly attached host. To avoid manual intervention protocols have been developed to allow for automatic exchange of configuration parameters ([46], [47], [48], [49]). Once all the devices on the distribution system have become aware of the new host and its location a static host will have full connectivity to the distribution system.

3.2.2.2. The Mobile Sender

The next distinction has to be made between the cases of a mobile sender and a mobile receiver. The solution for the first case is trivial and does not leave much space for sophisticated improvements since only the MH and the immediate BSs are involved and since no data packets have to be rerouted. If the sender is moving while sending, it will block the generation of new packets once a packet did not receive an immediate acknowledgment from its base station. As mentioned above this is standard behavior in wireless LANs in order to cope with the uncertain properties of the wireless medium. With the help of the reassociation mechanisms the mobile host will then associate with a new base station. The details of the reassociation procedure are not part of this thesis. Figure 21 shows the packet flow diagram that results for this case - after the reassociation is completed MAC packet transmission to the receiver continues. The packet flow shows a number of packets exchanged either across the wireless link or across the distribution medium. The packets used on the wireless medium are taken from the ones used in the IEEE 802.11 standard, since no changes should be imposed to the applied MAC layer interfaces. The applied frame types on the wireless link are:

- MAC data frame: MAC(sender, destination, real sender); the third address field is obsolete if identical with the sender field
- MAC acknowledgment frame: MAC_ACK(sender, destination)

The applied frame types on the distribution system are:

- Distribution system data: DSDAT(sender, receiver, real sender, real receiver)
- Distribution system acknowledgment: DSACK(sender, receiver) (this packet is optional; see Section 3.2.2.3. for a discussion on acknowledgments in the distribution system)

(other frame types will be used later on and will be explained where first used)

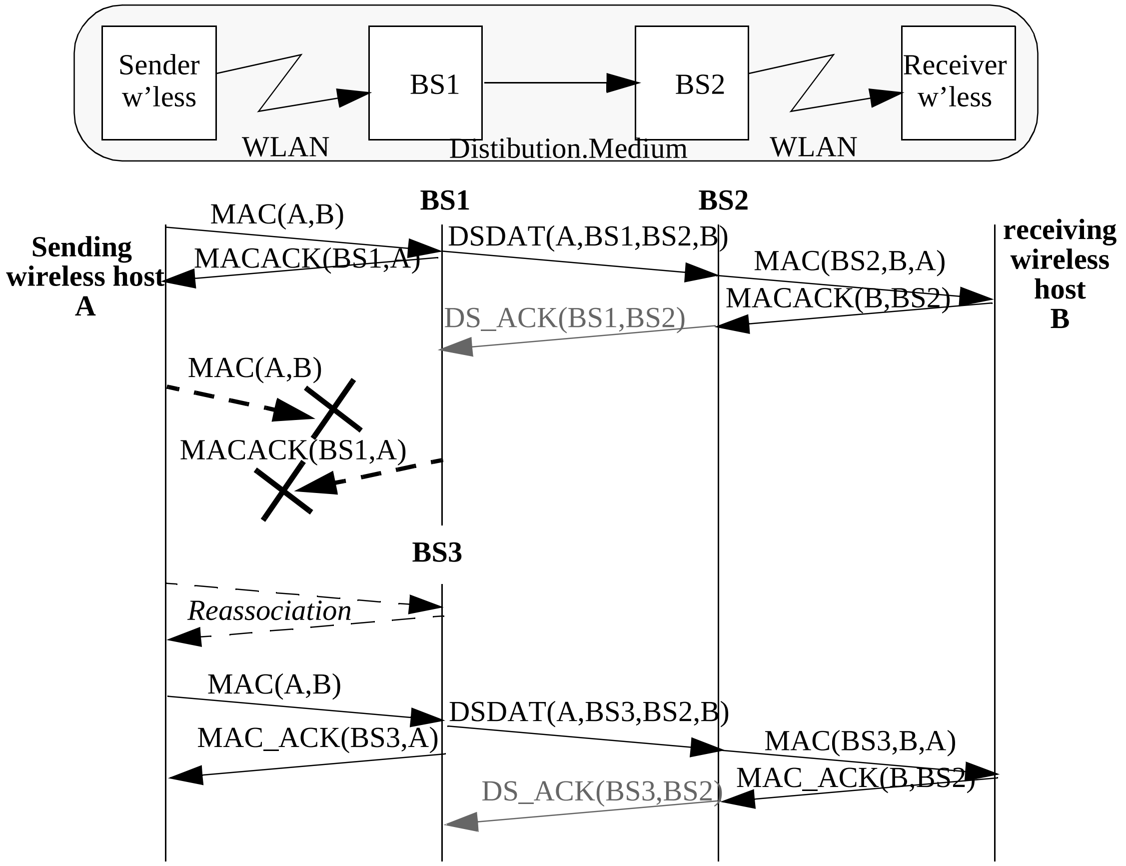


Figure 21:Packet Flow Diagram for a Mobile Sender and Fixed Receiver

After a successful transmission of a packet from A to B via BS1 and BS2, confirmed by the ACK that is sent upstream, the mobile sender A moves away. Depending on the moment when A moves either the next packet gets lost on its way to the base station BS1 or the acknowledgment misses its recipient. Which of those packets gets lost is of no importance. In any way the mobile sender will notice the unsuccessful transmission and will attempt to reassociate after a certain time-out (possibly after a number of retransmission attempts). As long as the sender does not send there is not much the distribution system can do to shorten the interruption duration. Only if the new base station BS3 completed the reassociation procedure with A, new packets may be transmitted. The old base station BS1 may eventually have to be informed of the new association. first however, and ask if the mobile host A really left the range of BS1 or if BS1 received the packet as well and will continue to serve A.

3.2.2.3. Acknowledgments in an E-WLAN

Two types of acknowledgment frames are applied in the packet flow in Figure 21: MAC acknowledgments and distribution system acknowledgment, the latter ones are marked as optional. With the introduction of the distribution system the interpretation of the MAC acknowledgments changes - the immediate acknowledgment no longer guarantees the successful transmission of the data frame to the final destination in the local network, the receiver or the gateway to other networks. The MAC layer acknowledgment only acknowledges the wireless portion of the intra-LAN path, the transmission from the wireless sender to the first base station, and the transmission from the second base station to the wireless receiver. This raises the question whether an E-WLAN-wide acknowledgment should be added, and if so, how it should be implemented. In the packet flow diagram in Figure 21 the MAC acknowledgment reaching the second base station from the receiving host triggers the generation of a distributions system acknowledgment sent to the sender-side base station. Upon reception at BS1 in Figure 21 the acknowledged reception at the final receiver is not used for anything else but to keep track of successful transmissions at the base station. The mobile host is kept uninformed. This is only one of a number of sensefull strategies of dealing with intra-LAN acknowledgments. Questions with acknowledgments are, that one has to decide whether

- the sender needs to receive a confirmation of the reception at the last host in the local net (i.e. that the old meaning of an acknowledgment in a LAN is reinstalled)
  - **and if so,** when and how should this confirmation be triggered
- if the wired (less error prone) part of the distribution system needs to be secured at all by immediate acknowledgments

Figure 22 presents a number of schemes, that reflect different answers to these questions. In scheme a) the sender receives two acknowledgments, one immediately to confirm the successful transmission over the wireless link, and a second one that acknowledges the reception of the packet at the receiver. This requires twice the transmission of an acknowledgment across the wireless link and twice acknowledgment processing at the sender. Also for the transmission of the second acknowledgment one cannot use the prioritized access mechanisms as provided for immediate acknowledgments e.g. in IEEE 802.11. The second approach b) delays the MAC acknowledgment until the successful transmission to the receiver can be confirmed, i.e. until the acknowledgments have been sent upstream over the distribution system. This scheme has the disadvantage that the MAC acknowledgment arrives much later than in the previous scheme with immediate MAC acknowledgment. The third strategy c) resigns from using any E-WLAN-wide acknowledgment, only the wireless links are secured by the immediate acknowledgments on the MAC layer, the sender does not receive notification of the successful reception at the receiver on MAC layer, the task of securing the transmission over the full path is left to higher layer protocols (e.g. TCP) or omitting it altogether[1]. The last strategy d) similarly uses the MAC acknowledgments only to secure the wireless transmission only but also applies an acknowledgment on the distribution medium, that is triggered upon reception of the MAC acknowledgment at the receiving base station and is sent to the sending base station.

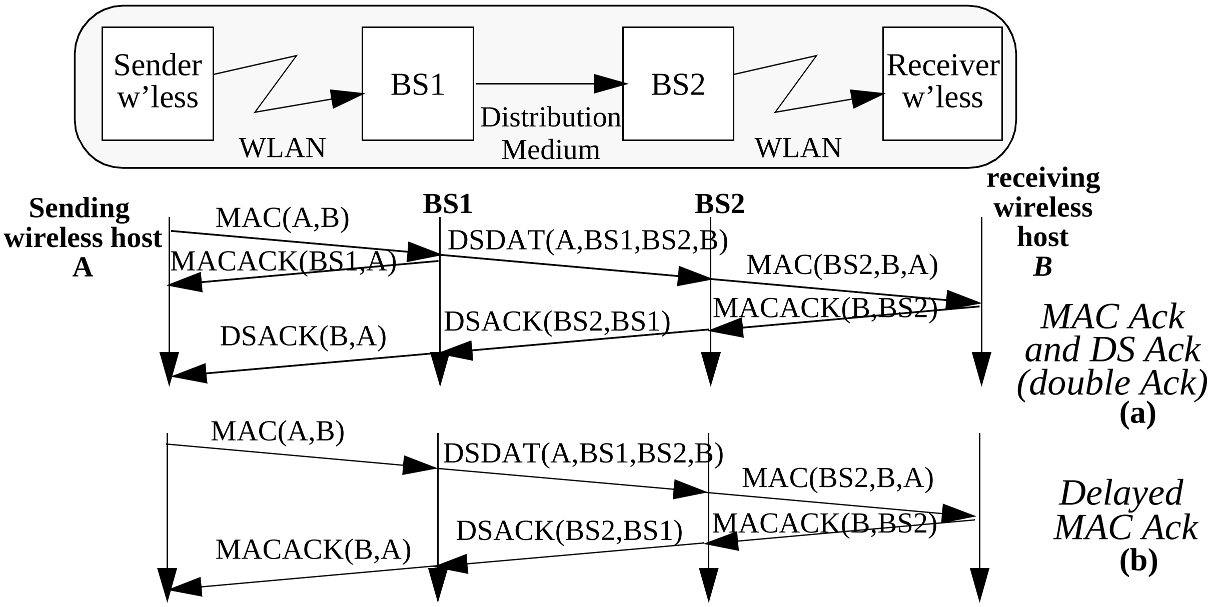


Figure 22:Acknowledgment across a Distribution System, part 1

---

1. A similar approach is described for Token Ring (IEEE 802.5) and FDDI. Both schemes contain a *Frame Copied-Bit* that is set by the receiving token ring adapter upon successful reception. But the advent of remote source-route bridging has meant separating LANs by great distances, thus creating possibly long delays. [50] describes a solution where a unique "local acknowledgment" capability in CISCO router/bridges solves this problem. When data is sent between two hosts separated by a WAN, the LLC2 session initiated by Host A terminates at the local router, which sends that host an acknowledgment within the prescribed time limit. To Host A, the acknowledgment seems to come from Host B. The data is then sent over the WAN to host B.

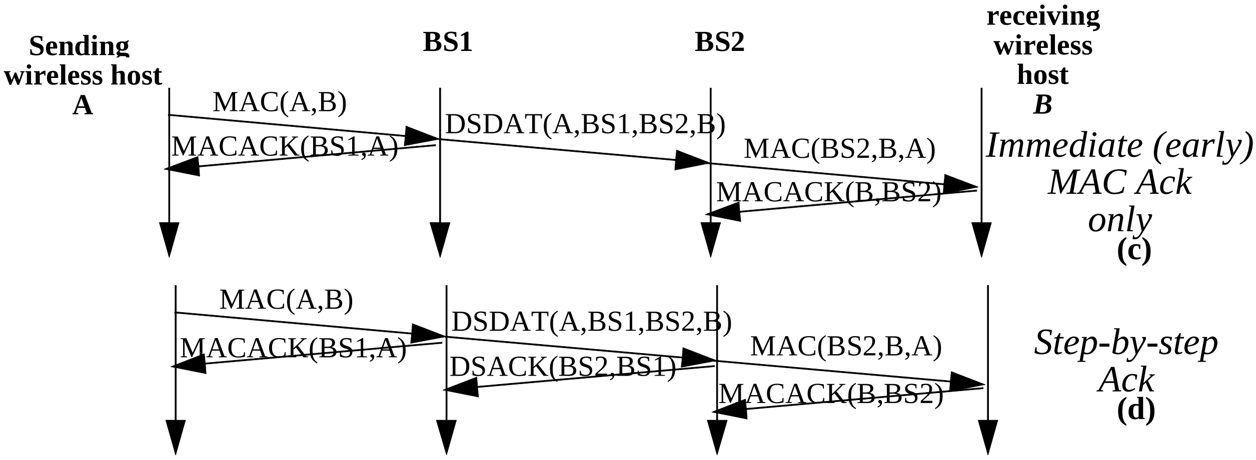


Figure 22: Acknowledgment across a Distribution System (cont.)

Which of the four strategies is applied has relevance on two issues: first - as already mentioned - on the meaning of each of the acknowledgments (syntax) and second on the time-out parameters that can be used if the acknowledgment is missing (effect). From the syntactical point of view applying the immediate acknowledgment only leaves the sender without any sure information of the reception of its packet at the receiver - even if it is located in the same LAN. However just as in Ethernet networks an acknowledgment mechanism on the less error-prune wired medium may not be necessary. The wireless sections of the path remain secured by the wireless immediate acknowledgments.

In any case a sender, that is not aware that the receiver may be located elsewhere, will surprisingly receive a MAC layer acknowledgment from another station other than the receiver - the data packet DATA(sender, receiver) will be acknowledged with an acknowledgment MACACK (base station, sender) or with a masked packet, that - even though issued from the base station - appears to be sent from the receiver.

3.2.2.4. The Mobile Receiver

Once the receiver becomes mobile there are a number of new problems to solve. First one needs to find the correct location of the MH again at the **present** moment i.e. one has to keep the location table correct. This updating process needs to be initiated somehow and carried out as fast as possible. Furthermore one has to find a way how to deal with packets already on their way - should they be dropped or forwarded. How should the route to the present receiver-location be adapted? How does the MH discover the existence of a BS or how the BS the MH? Solutions for these issues will be presented in Section 4.

3.2.3.Packet transport, Addressing Format

Once the correct location of the receiver is known, the data needs to be transported to this location. To do this an addressing scheme is needed that reflects the location of the host within the LAN. This information is not provided by the two addressing spaces available on MAC and network layer MAC addresses and IP addresses. Possible solutions can be either introducing a new address space compared to the two mentioned before or using the existing addresses with new interpretation.

3.2.3.1. New Address Space

As an example for the introduction of a new address space - compared to the existing MAC addresses and network addresses - can be found in the Hiperlan draft standard [4], where a certain identifier (Hiperlan-ID) to each Hiperlan is assigned to distinguish it from other Hiperlans.

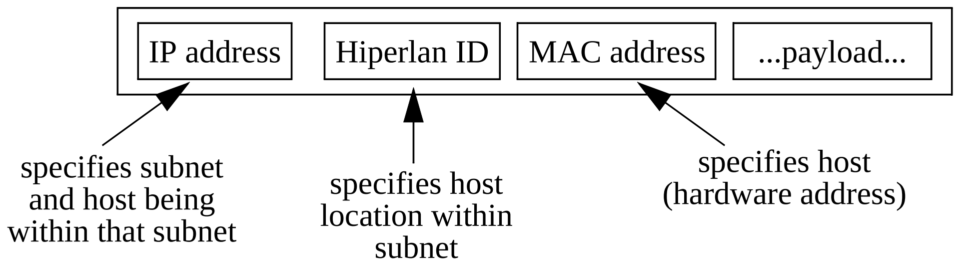
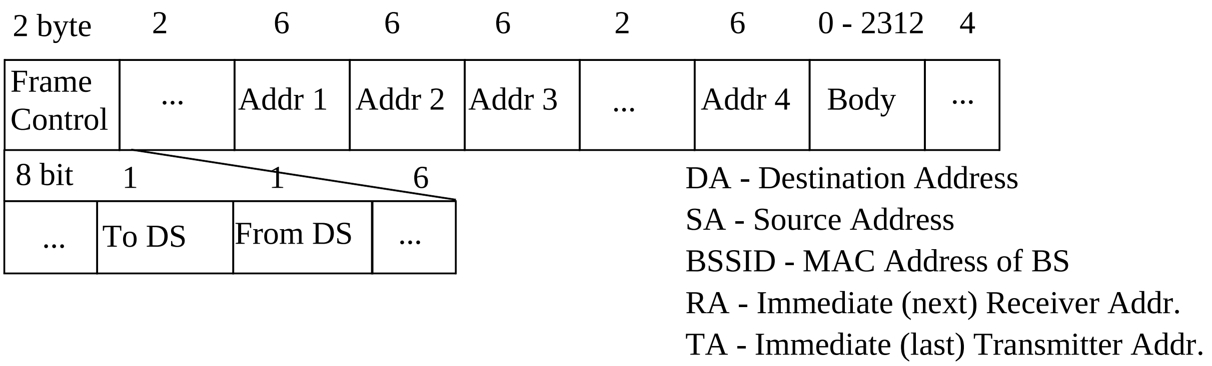


Figure 23:Application of New Address Space in the Hierarchy of Addresses

A new Hiperlan can be created by choosing an identifier which is not in current use in the communication environment. An already existing Hiperlan can be joined by just using its identifier, it is left by refraining from using its identifier, it is destroyed when no more entity uses its identifier. The Hiperlan-ID is used to address packets ***to*** *a destination host* ***in*** *a certain hiperlan*. I.e. there are two destination address fields in the packet - Hiperlan-ID specifying the location and MAC address specifying the host. One gets a new hierarchical layer of addressing between the subnet-part in the IP address and the host address (either in the host-dependent pat of the IP address or in the MAC address). The introduction of the new address space allows the parallel use of several distinct but overlapping Hiperlans (e.g. for a redundant installation for backup purposes or for bandwidth increase).

3.2.3.2. Multiple Address Fields

The other possibility opposed to introducing a new address space is to use the existing addresses with new interpretations of the address. IEEE 802.11 [3] has defined its packets in a way that such a scheme is implied: the packets contain 4 address fields and a field indicating the meaning of the address fields(Figure 24). To fill in the correct values the originator either has to leave fields empty or has to know about the location of the recipient, at least to the point whether it is in the same cell or in another cell of the E-WLAN. This information cannot be supplied by common address resolution mechanisms.



To DS	Frm DS	Addr 1	Addr 2	Addr 3	Addr 4
0	0	DA	SA	BSSID	N/A
0	1	DA	BSSID	SA	N/A
1	0	BSSID	SA	DA	N/A
1	1	RA	TA	DA	SA

Figure 24:IEEE 802.11 Frame Format and Address Field Definition

In both cases each packet has to contain 4 full length address fields increasing the overhead in each packet. If the second scheme is applied at each intermediate device the packet header has to be rewritten causing a higher processing delay compared to the first scheme where the packet may be forwarded unchanged. However if only available address spaces are used the integration into existing network architecture is less complicated.

3.3. Components of a Distribution System

In the following sections I will discuss the issues concerning the different components of the distribution system, the mobile host, the base station, the distribution media, the portal and the hosts directly attached to the DS.

3.3.1.Mobile Host

One issue in regarding the design of the MH is the question if active participation can be expected in the handover detection, handover initiation, handover processing, reregistration, or if the limited resources of the mobile are too scarce to allow additional activity. Portables should allow a working time of 4-6 hours without the need of recharge. Unfortunately, if modern portable computers are connected to contemporary WLANs the longevity of the battery can be significantly reduced. In [53] it has been pointed out, that contemporary wireless network interface cards (with 1 Mbit/s bit rate) can take 12 times more power than a standard 10 Mbit/s ethernet card. Thus it should not be surprising that longevity reduction in the range of 60% has been reported for portables in [54]. Depending on the way of usage, during our lab work we have observed a dramatic decrease of time in action from 3 hours to 45 minutes with a laptop using a WLAN PCMCIA interface card.

The second main issue that has to be faced regarding the mobile host is that one cannot expect to impose changes in the MAC protocol, MAC layer interface or in the MAC layer specific hardware. Therefore the application of new MAC-packet formats should be avoided as far as possible. Every added feature must be able to gain all the support needed from the MAC layer interface as sold.

3.3.2.Base Station

The base station is an internetworking device like bridges or routers. It is equipped with two interfaces - one to the wireless side, one to the wired side and has to transfer data from one input to the other based on. As an internetworking device it has similar design problems like those well known devices such as the trade off between buffer size and mean processing delay, the problem of buffering between two mediums that offer different transmission bandwidth and the limited capacity on the internal bus. To enable the base station to deliver queued data to the mobile hosts it might need to have prioritized access to the wireless link. The asymmetric distribution of load between the separate hosts and the base station that has to serve all hosts in its cell could otherwise cause large processing delays and buffering overflows at the BS. With its prominent role as the link between the wired and the wireless part of the E-WLAN and its easy access to resources like electric power the BS is a perfect candidate to accommodate for most of the functionality for the distribution system. If one assumes a passive MH than one will have to put most of the load on the BS to provide most of the intelligence of the DS. The tasks it will have to fulfill besides inter-cell networking and delivering data from the DS to cell, are readdressing, location management, handover management.

3.3.3.Distribution Media

An interesting question with wide effect on the overall system performance is the choice of the medium that shall be applied in the wired part of the distribution system. Any technology over any medium - from copper to fiber to wireless - using any access scheme may be used. Preferably of course one will chose a technology that is readily available on the market due to its lower prices. One can imagine a distribution system via classical or high speed LANs (Ethernet, FastEthernet), a distribution system via ATM (with support for mobile VCs), or a distribution system via optical networks. An ‘of-the-shelf’-technology also helps to keep the solutions easily compatible with current technologies. Therefore the common LAN technologies are the candidates to be looked at for suitability. Relevant issues are hereby logical topology, capacity, delay characteristics, features of the access scheme and scalability. Regarding the logical topology one may use bus, ring, star or mesh. Where a bus might offer higher robustness than a ring or a star, it is far inferior to the star in terms of scalability. A bus or a ring with easy broadcasting facility will be preferable if self-learning of location information is applied. Regarding the capacity that is available one might prefer a solution that offers *fixed amount of bandwidth to each cell* (like switched ethernet) compared to a solution where a *fixed amount of bandwidth has to be shared between all cells*. If the access scheme is random based with its statistical, non-deterministic behavior it can cause unevenly distributed delay - potentially of unlimited length. Therefore one might want to choose an access scheme that provides resource reservation.

The distribution media offers certain characteristics like delay, bandwidth/channel capacity and maximal supported distance. Those characteristics need to be tuned in order to match the characteristics of the connected networks, i.e. it should be able to accommodate all traffic that intends to pass across the distribution media - be it from outside network or from the connected cells. It should not add unpredictable delay, and, of course, as little as possible.

3.3.4.Portal

Even though the portal looks like an ideal candidate for centrally located functionality it is unusable for this purpose since E-WLANs like every LAN may well be established without inter-connection to an internetwork. In such an installation the portal will not be present. - or would have to be separately installed - but still the mobile users will want connectivity between the cells of the E-WLAN. Therefore the distribution scheme may not rely on portal but has to tolerate its presence. Just like the base station it is an internetworking device like bridges or routers equipped with two interfaces. It has to be dimensioned carefully so that the limited throughput in packet/second and the limited buffer space suit the requirements of the network. It differs however from the base station in a way that it has to deal with a special kind of traffic - traffic from outside networks originating from hosts that may be communicating from far away. It may have to handle long distance communications with its distinct requirements compared to local area packet exchange. It may also be an appropriate device to incorporate the mobility management functionality on higher protocol layers (e.g. operating as the home location register, HLR, or visited location register, VLR, for MobileIP on network layer) to handle inter-LAN mobility, since this functionality is **only** needed when a portal in fact is present in the system.

3.3.5.Hosts on the DS - Printer/Server/...

One would want to use off-the shelf devices for print servers, file server and other directly attached hosts therefore one has to accept distribution-system-blindness by those hosts i.e. they do not participate actively and are not aware of the different environment. However this lack of participation requires to provide for the needed functionality elsewhere. A solution might be to connect the network access of the directly attached devices to a small transceiver-like box, that provides the necessary additional functionality.

3.4. QoS Issues in a Distribution System

If an E-WLAN is supposed to be used in place of (or parallel to) a wired LAN, it is likely that support for quality of service guarantees will be expected to a certain degree[1]. According to [55] support for QoS consists of several sub tasks:

- QoS identification
- QoS specification
- QoS mapping
- QoS communication
- QoS negotiation
- QoS enforcement

In the setup chosen here most of those subtasks are out of the scope - they have to be implemented on an end-to-end basis or with the help of a resource reservation agent, whose design is widely independent of the system used here. In the chosen setup the task to fulfill is to maintain the QoS from the wired part of the internetwork also across the E-WLAN part of the path. This again first has to be looked at for the static case, i.e. how can the desired QoS be carried over the full extend of the E-WLAN.

The path within the E-WLAN consists of 2 hops (or 3 if both sender and receiver are located within the E-WLAN), one (or two) wireless hop to the base station and a hop across the distribution medium. First this subsection will present a look on the abilities of recently completed standards for WLANs - the HIPERLAN standard developed by ETSI and the 802.11 standard developed by the IEEE - that are both representing possible candidates for the wireless hops. The issues in the remaining devices of the distribution system - the base station and the distribution medium - arising with QoS provision will then be pointed out. The last subsection contains a discussion of the problem of maintaining QoS in the event of a handover.

3.4.1.QoS in a E-WLAN, Static Case - Quality of Service in WLANs and the DS

Although the HIPERLAN claims to support time bounded services it does not provide any services that guarantee Quality of Service requirements. The idea behind the concept chosen for HIPERLAN is that the LAN should transmit a time bounded packet first before a packet which is not time constraint and should transmit a packet with a short deadline before a packet with a longer deadline. To realize this concept the channel access mechanism provides non pre-emptive priorities. Any node automatically defers before any other node about to transmit a packet with higher

---

1. Wired LANs like the widely used Ethernet do not provide support for QoS so far, mainly for the reason that such demand was un-thought of at the time of development. However the demand for QoS cannot be neglected for **new** developments

priority. Therefore one can say that the HIPERLAN allows to distinguish between traffic classes but it does not support the allocation of a fixed portion of bandwidth nor any other QoS parameters. Thus, HIPERLAN is still just a best effort network and as such not suitable to extend ATM networks or other QoS-supporting networks from the QoS point of view.

Simulations of a HIPERLAN network with 6 sender, that attempt to transmit time bounded traffic at a fixed rate of 100.000 byte/s are shown in the below figure. They send with high user priority whereas the parallel increased background traffic is transmitted with low user priority. The effect of both traffic classes against each other is shown in Figure 25. The throughput and delay of the time bounded traffic remains unaffected from the increasing background traffic due to its higher priority. This good behavior for the traffic classes against each other however can be damaged simply if too many stations are attempting to send with high user priority. There is no mechanism to protect the QoS from competition with too many equally prioritized stations. New stations coming in into a cell after a handover may damage everyone’s QoS.

An interesting aspect can be seen in the right graph of Figure 25 showing the corresponding delay: Hiperlan does not deliver any packets that have been aged beyond a certain limit. Those old packets are dropped to protect the network resources from transmitting unnecessary information. Under high load one can assume that many packets that did not reach the destination on time are dropped. Since in the simulation only packets are counted that actually have reached their final destination all those dropped packets are not reflected in the delay curve. The delay value is mostly limited by the maximal life time as set in the simulations (here 100 ms for time bounded traffic, 1000 ms for background traffic).

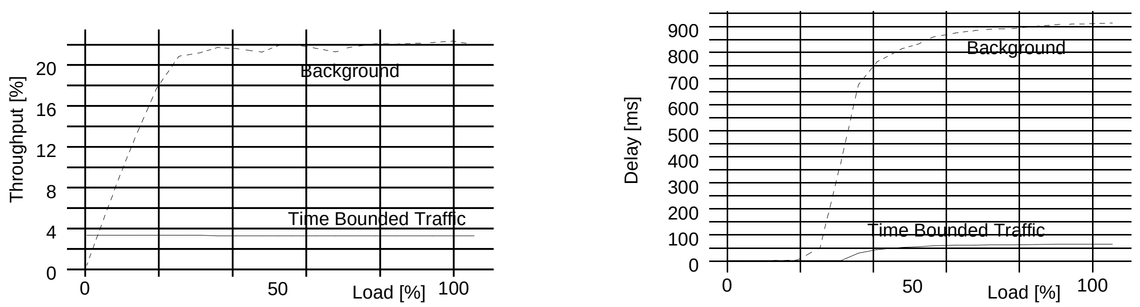


Figure 25:Throughput and Delay with Time Bounded Traffic Hiperlan

The IEEE 802.11 draft standard offers support for time bounded services by integrating the before mentioned point coordinated mode in which a centralized controller gains control over the networks resources and as such is able to guarantee a fixed portion of these resources to stations requesting it. For brevity reasons only the results of several simulations regarding the performance of this centralized mode will be described. Unlike in Hiperlan any background traffic does not have any influence on the throughput time bounded traffic due to the guaranteed reserved bandwidth. However once more time bounded bandwidth is requested than is available, all bandwidth requests may end up unsatisfied. This depends on the strategy that is applied to satisfy requests: reduce the available bandwidth for everyone or refuse new additions to the polling list. This has not been specified by 802.11. Other unspecified issues in applying the concept of centralized polling are e.g. what strategy is used to register inside the WLAN for the services guaranteed, how the WLAN QoS is mapped to the QoS on other links of the end-to-end-connection, how to adapt to QoS-destroying configurations like hidden terminals or large fluctuations of stations in one cell etc. Another uncertainty comes with the proportion of each part within the superframe. If the length of the DCF part is large compared to the PCF part, the advantage of the PCF does not become apparent. However if the PCF part is large compared to the DCF part, all traffic flowing across the access point (i.e. all traffic going into or coming from the attached infrastructure) has an unfair advantage to gain access compared to direct traffic between mobile hosts, reducing the available bandwidth for the intracellular traffic.

To sum up, the HIPERLAN standard does not allow to guarantee a fixed portion of the available bandwidth to a host requiring such service, nor does it give delay guarantees. It sacrifices this explicit support for QoS in favor of an implicit mechanism that puts preference on performing as good as possible under the current condition for each priority class. This implicit QoS support may prove useful however for the wireless environment with its unpredictable changes in the communication path due to external influences (e.g. distortion), but it is more vulnerable to QoS-threatening configuration changes such as too many users requesting transmission in the same high service class. IEEE 802.11 - with its point coordinated mode - provides mechanisms for bandwidth reservation, however several questions regarding its applicability remain unanswered so far.

If the WLANs are assumed to be capable to support QoS, one has to look at the remaining devices along the path. Those are the base station (and possibly the portal, however with the same issues concerning its design as the base station) and the distribution medium. The distribution medium should simply not add much influence on the experienced QoS characteristics as there are mainly delay and delay jitter, and should offer sufficient bandwidth to not become a bottleneck.

The base stations have more complicated requirements imposed on them. First - just like the distribution medium - they should not introduce a significant portion of delay and delay jitter to the data passing through it in both directions. To do so it has to be capable to process the incoming packets at line speed. However since the bandwidth available on both sides of the base stations is likely to differ it will have to provide for buffer space to allow temporary packet buffering, if the incoming data cannot be sent out again at the same speed, which will cause delay increase.

In order to coordinate the QoS support over all parts of the distribution system some part of the resource control has to be concentrated at a resource reservation agent. Therefore the base station has to either pass over the control of the resources it is controlling - bandwidth on the wireless side, buffer space in the base station and access priority on both sides of the base station - or enforce these reservations itself based on the requirements received from the agent. For an integrated approach one also has to map the QoS classes between the different links and layers - the WMAC-QoS parameters, DS-QoS parameters and network-layer-QoS. This task is left open for future research.

3.4.2.QoS in the Case of Handover

The following section will contain a qualitative discussion of QoS Issues in a Distribution System in the event of handover. Can the system keep the guaranteed QoS in the case of handover for the ongoing communication, and are special access control or performance control mechanisms needed to protect the QoS of communications of hosts that already present in that cell.

As discussed before a handover will cause an interruption of a certain duration and therefore cause a certain packet loss. If this is very much unwelcome for the application - if the flow is delay- and loss-sensitive to a high degree. As the user moves, it may be the case that no resources for it will be available in the new cell. How can support be granted for a guaranteed and predictive service. How can it be avoided that an ongoing communication has to be dropped - which is likely to be considered more harmful that to be blocked before the communication is started. Some calls however must be degraded or dropped (though not necessarily those of the host that just entered the cell).

In mobile telephone networks the approach usually implemented is the reservation of a number of channels (“guard channels”) dedicated for users moving into the cell. each requested entity has the same bandwidth requirement. In a multimedia internet environment usually flows of different dimension and requirements are found - data, voice, video with or without color, with small, medium, large resolution.

If this full pre-reservation strategy is followed also in the described environment one would have to reserve network resources in all the cells the mobile host can move during the life of the session. By doing this one can be sure that resources will be available everywhere as the mobile moves to other cells. However this will lead to a very low network utilization since bandwidth that was reserved in cells in which the mobile host is not currently there are wasted.

A different approach would be to reserve and use the resources only as the mobile host enters a new cell. By doing this one can have high network utilization since all the resources reserved are being used for active hosts. However it may be the case that the mobile host will enter a congested cell where resources for a new user are not available.

If a large number of mobile users are present one may attempt to reserve only a certain percentage of the required bandwidth in the neighboring cells and to benefit from statistical multiplexing in a way that not all hosts will eventually need their pre-reservation share and may therefore provide for sufficient pre-reservation for those hosts actually moving into a certain cell.

Any solution will have to closely cooperate with the access control mechanism. To minimize disruption to the user, one may want to make the possibility of degradation apparent to the user when they select the necessary QoS for their calls. The user should then be able to group calls and specify degradation paths such that the resource control mechanism can use a ``QoS calculus'' over the calls in a cell to decide which calls should be degraded or dropped. This necessitates a dynamic aspect to call QoS management and thus additional signalling. If the flow is likely not find sufficient resources over its lifetime one may

- turn down flow
- give guarantees for static host, no guarantee if becomes mobile
- propose flow adaptation to mobile to available resources
- search for free resources (ask other mobiles for adaptation)
- propose delayed admission

The previous subsection has qualitatively discussed the issues arising with the integration of QoS support in a E-WLAN. A number of questions touched here still need to be solved. This is left for further studies.

Chapter 4

Designs & Specifications

4. Designs & Specifications

As discussed before, solving the case for the mobile receiver is difficult and is the most complex part of the problem. The reason for this is the chosen connectionless environment that does not allow the immediate notification of all current communication partners - this information is not available - and the assumed non-participation of the mobile receiver, to reduce complexity, processing and power consummation in the mobile station.

In this chapter a number of different solutions will be presented for this case, each with a different priority in the design and therefore with different features and different performance characteristics as shown later on. The following section starts with general considerations and develops a systematic classification of solutions. Following a presentation of two types of periodic schemes the issue of how reactive solutions may be triggered is discussed. The chapter then develops a solution that is based on reactive procedures after the handover event. A specification for each of the solutions is presented in the respective subsections.

4.1. General

As already mentioned it is assumed that all hosts have passed a number of initialization steps when they first enter the WLAN. Part of these procedures is the registration of the mobile host as member of the local network with an identification, authorization and registration instance. At this moment there is either no ongoing communication happening or a inter-network handover (handled on higher layers, see Section 2.3.) with its longer interruptions to the ongoing communication is occurring. For the distribution system it is not advised to attempt to minimize this period since those efforts may interfere with procedures on other protocol layers and may therefore proof to be counter-effective. Registration ensures that the base station has the knowledge of its currently served MHs, however with those hosts becoming mobile the knowledge may be imprecise, incorrect, or inconsistent. It may cause problems if this inaccurate information is passed on to other stations as a reply to searches.

The method applied for this initial registration procedure - while itself being outside of the scope of this thesis - has some impact on the solutions discussed later on for the distribution system. Generally for an initial registration procedure a *registration protocol* is needed, a *registration database* in a *registration server* and a scheme to *discover the registration server*. The registration protocol supporting the scheme to guarantee authorization and secure identification is necessary for any kind of network with mobile hosts, no matter if with or without a distribution system of the chosen type. Also the question how the new mobile host will make the first contact to the registration server is not a distribution system specific issue. Solutions here may be periodic server announcements or a request sent to a well known address (anycast address).

However the solution for the registration database and the registration server has impact on the schemes developed in this thesis since it is this database that needs to be kept up-to-date in the mobile environment. The primary distinction in this context has to be made between a centrally located and a distributed database. A distributed database, in addition to the effort to get the current information, requires effort to keep the information in the different elements consistent, but it scales much better in terms of number of supported hosts and processing requirements. A central database forms a single-point-of-failure, if not backed up with a redundant mirror. A central database may be reached by updates faster than a distributed one however in the latter case the information is likely to be available nearer (and thus earlier following a request) to the place where it is needed. If one look at the targeted environment and estimate the expected number of users and the expected number of requests, the currently available database technologies will be able to support the dimensions of this database problem. Experiences gained from core Internet routers show, that databases nowadays are able to support a database of many more entries than one might expect in this environment (see Section 1.6.) while still keeping very fast response times. So apart from the concern that a single central database is a single point-of-failure in the system and that one may not want to have a central device at all a central database would serve well in this case.

4.2. Systematic Classification of Schemes

As discussed in Section 3.2.2.1., Section 3.2.2.2. and Section 3.2.2.4., the problem to solve is the (static and dynamic) location management with mobile receivers. No active participation of the mobile receiver in providing the information of its location can be assumed - the mobile receiver is not aware who might need this information at this moment or in the near future. The mobile receiver will only carry out the association/reassociation procedures, that are part of the WLAN protocols like IEEE 802.11[3]. The location of the mobile receiver is therefore initially only known at its base station, where it associated. The problem that arises is: how can a sender find out about the current location of a receiver, that it may want to start communicating with.

Basically this problem of location management is a question of data distribution. The two elements of such strategies are *Update* and *Search*. During the update procedure the current location information of the mobile station is distributed into the network up to a certain degree, during the search procedure this information is searched. Once the location information is available where needed, messages can be delivered to the host by making use of commonly applied packet forwarding mechanisms. The two extreme ends of this trade-off are either *inform-all-over-the-network* or *search-all-over-the-network*. In the first case the location information is distributed to every routing station in the network, causing a large amount of traffic. This, in turn, will result in a minimal effort to lookup the information, once needed, since it is available everywhere. In the latter case, no information about the current location is sent anywhere, placing the load of the locating task completely on the search procedure. Once needed, the location information has to be searched all over the network, since it has not been distributed at all. Obviously both strategies do not scale very well for larger networks, the first scenario produces large amounts of update traffic, the latter one large amounts of search traffic, however these two poles limit the ‘playground‘ for all location management schemes. Since the chosen architecture explicitly addresses a network spanning only a local area this scalability concern does not fall into account.

These considerations lead to a classification for approaches that may be applied for the problem under consideration, as shown in Figure 26.

- A “periodic table” approach, where the base stations periodically exchange the lists of their currently served mobile hosts. Packets that are send to a wrong destination due to old location information are merely dropped. A full discussion and specification of this approach can be found in Section 4.3.. This can be seen as the most simple approach. However it causes high control (update) traffic, and this overhead occurs even if no communication takes place and even if the corresponding hosts remain static at one location.

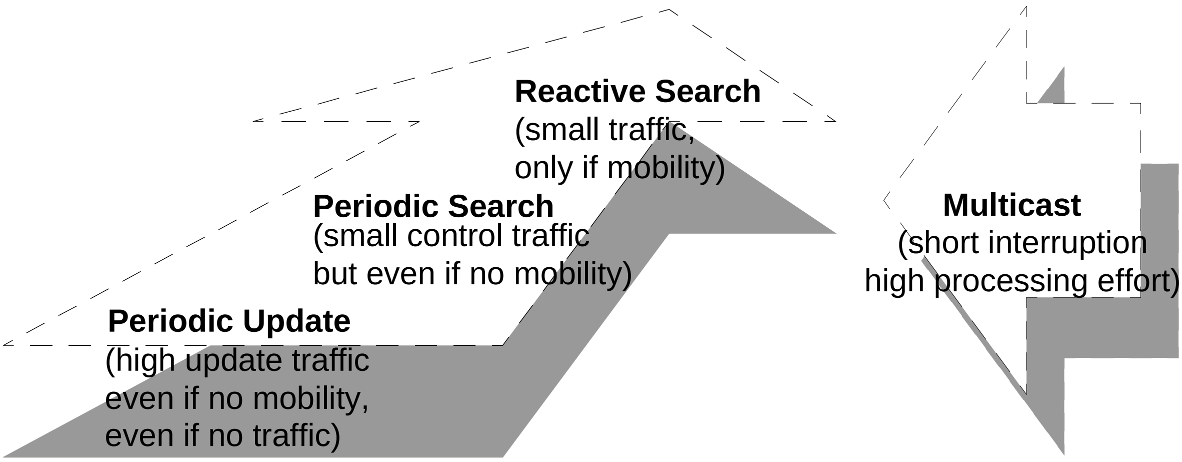


Figure 26:Classification of Schemes

- A “periodic-search” approach, where the base stations search for the location of the receiving host upon initial communication request and periodically re-search to keep up-to-date information. This approach is specified in Section 4.4.. It solves the first of the two problems of the periodic update approach, since it does not generate control traffic if no communication takes place, however still its overhead is independent if the receiver is static or mobile.

- A “reactive-search” approach, where the sender itself initiates a search process for the location of the receiver. The search is re-initiated if transmission failures occur, indicating a handover. In order to be able to react on a handover, the sender has detect the occurrence of the handover, at best as soon as possible. Since the system is lacking active handover notifications it has to rely on timer-based detection schemes, which is non-trivial to implement as is explained in Section 4.5.. However if the detection can be done successfully, the number of search processes is reduced to the number of absolutely necessary ones. In my model the handover detection is done by a negative acknowledgment (NACK), that is sent upstream from the old base station if transmission to the - now gone - receiver failed. The approach is specified and discussed in Section 4.6.

The last approach, that I will present is outside of the *search-or-update* classification scheme, since it is not designed to provide location management in the chosen environment. Its target is to minimize the interruption that is caused by the handover event, i.e. to optimize the behavior in the event of dynamic location changes.

- The “multicast” approach adds functionality on top of the previously descried approaches. In my simulation I have built the multicast scheme on top of the periodic search approach, however any of the other two schemes could be chosen as well, which however would not yield significantly different results. All packets are transmitted to the base station serving the receiver as well as to a group of neighboring base stations, where the data will be stored for the case that the receiver will move to this cell in the near future. Section 4.7. presents the specification of this scheme.

4.3. Periodic Table Scheme

As discussed before in Section 2. many types of devices (e.g. bridges, LAN-switches, repeater) have been developed to *interconnect* classical wired subnets without offering any support for *mobility management*. However in many application scenarios of the chosen architecture this might not be required - plain interconnection with a certain, reasonably fast configuration updating scheme might be sufficient in a quasi-static environment. Therefore initially a ‘periodic table approach’ will be developed, that operates as follows: the wireless pico-cells are connected via common bridges (or other interconnection-offering hubs) to a fixed backbone network. The periodic table approach only provides basic functionality, i.e. interconnection for static hosts and very limited service for hosts that become mobile and change cells. Connectivity is only granted after the update of the connectivity tables has taken place between the bridges. The DS would not bother to deal with buffering or rerouting - packets on their way in the system at handover time will be dropped. As a result the system requires very little administrative complexity. A mobile receiver that moves to a new BS has to notice the handover himself, eventually with the help of periodic BS advertisements and initiate the reregistration. At this time it will not yet be able to receive again. It will have to wait for the exchange of routing tables, that are periodically distributed among the base stations. Once the tables have been exchanged the new location of the mobile receiver is learned by all participating base stations. However all packets that were already on the way or transmitted during the handover phase will be lost. One has to rely on higher layer retransmissions (e.g. TCP). Its design allows for high access security and offers easy administration since explicit reregistration is required but it does not care about the length of the service interruption.

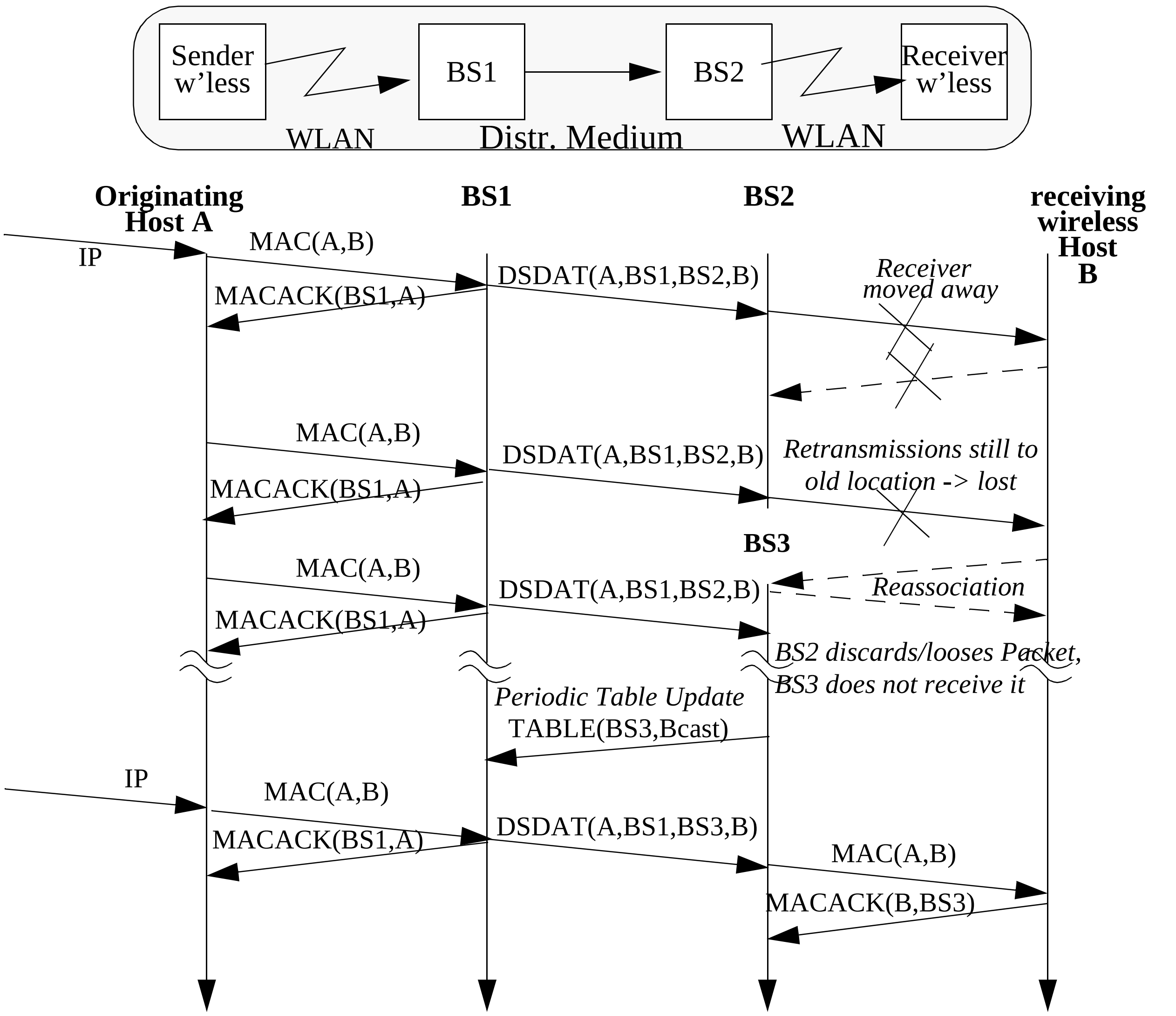


Figure 27:Periodic Table Scheme, Packet Flow

Figure 27 shows the packet flow in the periodic table scheme. The MAC frame, sent out from station A to station B is immediately acknowledged by the base station that also passes the frame on across the distribution medium to the correct base station. The base station takes this information about the correct receiving base station from its connectivity table, that has been built up with the help of previously received table exchange transmissions from the other base stations. Since however the receiving mobile host moved out of reach from that base station the packet is subsequently dropped/lost at the - now no longer current - base station (it does not make a difference if the packet never reached the mobile host or if it got through to the mobile host successfully before it left the range of its base station but the acknowledgment did not reach the base station, since in both cases the transmission remains incomplete) Further packets, that are sent out by the sender-side base station with the address of the old base station, unaware of the handover are lost as well. This happens even **after** the reregistration of the receiver since the base station, that grabs the packets from the sender, still has no updated information. Only after the new base station, that now serves the receiver, broadcasts its new table with the new entry for the mobile receiver, which is done periodically, the receiver can again be reached.

The following frames have to be used.

- MAC data frame: MAC(sender, receiver)
- MAC acknowledgment frame: MACACK(sender, receiver)
- Distribution system data: DSDAT(sender, receiver, real sender, real receiver)

The first two are frames as applied in the WLANs, the third frame used to transfer data over the distribution system, is a MAC data frame only extended by two further address fields defining the path across the distribution system. In order to be able to transmit the tables over the distribution system a fourth packet format is introduced, that only requires one address field, since it is broadcast:

- Table distribution packet: TABLE (sender)

The state diagram for the mobile host (Figure 28) for the periodic update scheme shows two states on the sender side of the mobile host (Idle and Wait_ACK) and one state (Idle) on the receiver side. In the sender path, IP packets may arrive from higher layers to be sent to other hosts or acknowledgment packets may arrive over the wireless link. Once a MAC packet has been generated the sender has to enter the state Wait_ACK where it remains until the transmission is acknowledged or until to many (timer based) retransmissions have been attempted. While being in this state further transmissions are blocked. In the receiver path only MAC packets are received, acknowledged and passed on to the higher layers.

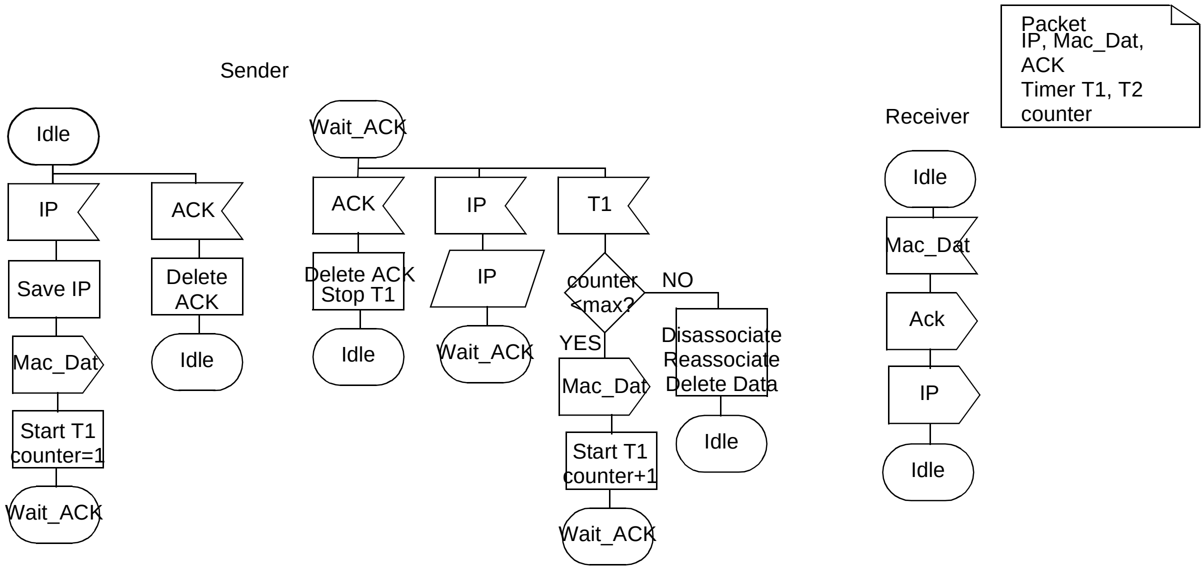


Figure 28:State Diagram Periodic Update Scheme, Mobile Station

As can be seen in Figure 28 no special mobility related functions are implemented in the mobile station. They have to be located in the base station.

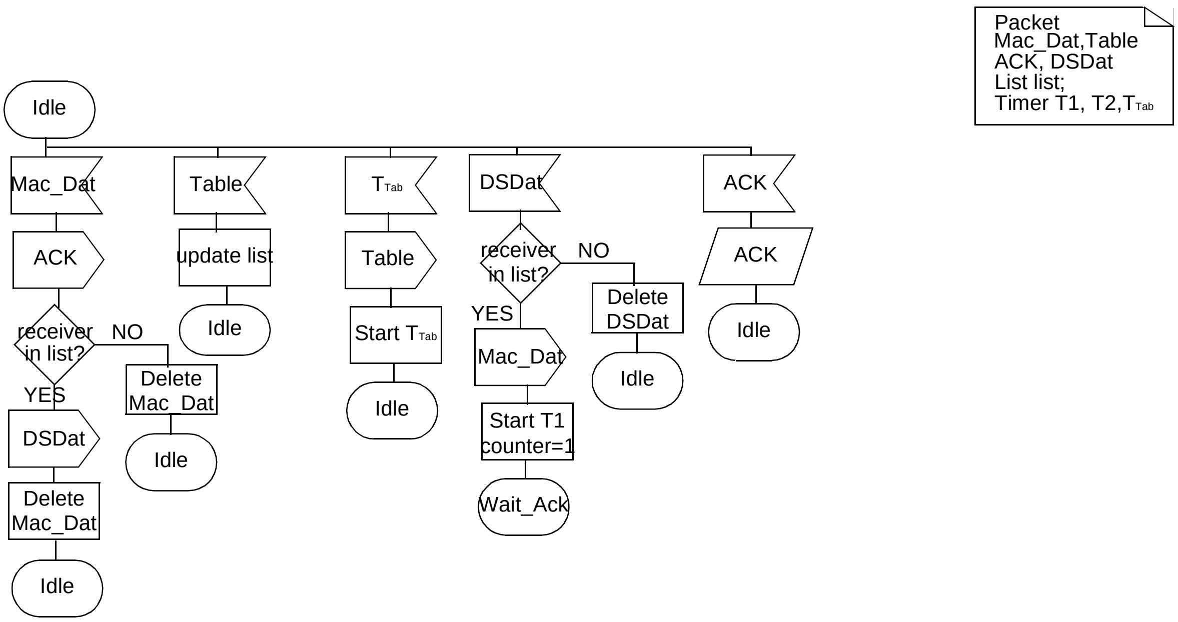


Figure 29:State Diagram Periodic Update Scheme, Base Station, part 1

The state diagram for the base station (Figure 29) shows that in addition to the packet forwarding mechanisms, that take packets from the wireless side of the base station onto the distribution system and vice-versa, periodically tables are sent out containing the list of mobile station, that are currently served by the base station and that are used upon reception to update the routing list. The parameter of the table cycle timer set in all base stations obviously is the most important parameter of this scheme.

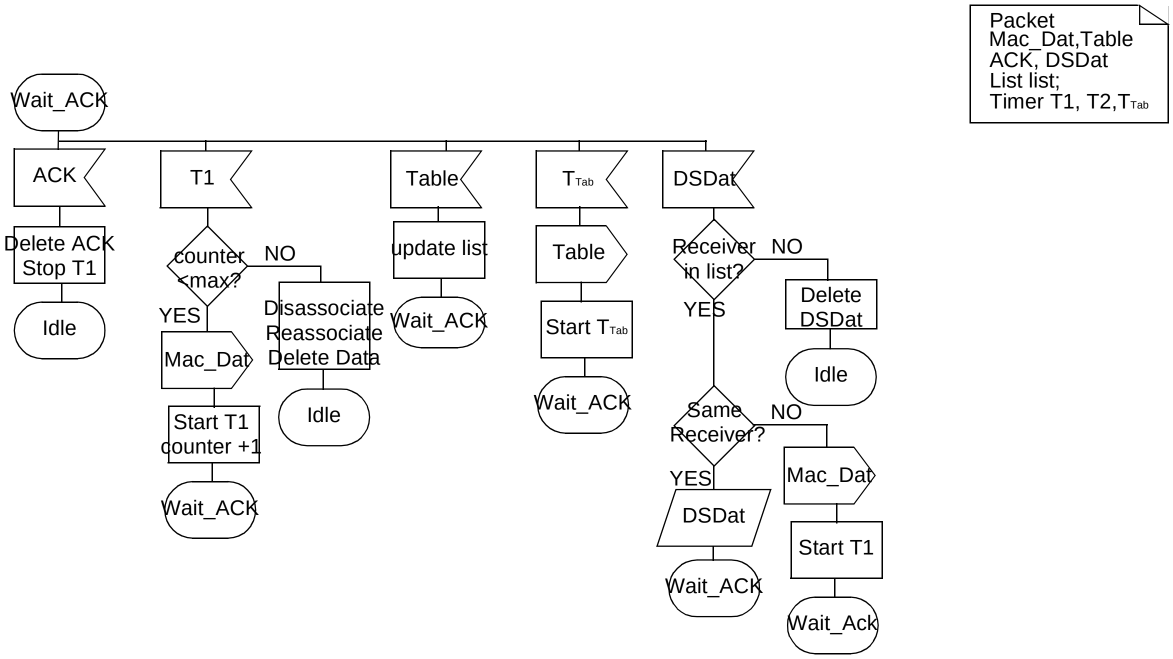


Figure 29: State Diagram Periodic Update Scheme, Base Station, part 2

In the periodic table scheme the shortest possible interruption cannot be shorter than until the next connectivity table update cycle is completed. If common bridges are applied the update may not occur before a couple of seconds have passed. Before this update has taken place each attempt to reach the mobile receiver will be unsuccessful. However since the ongoing connection will be terminated anyway this interruption is not likely to be a serious limitation. From the bandwidth efficiency point-of-view the packets, that will be dropped during the cell change are undesirable, since the scarce wireless bandwidth will suffer from additional overhead.

4.4. Periodic Search Scheme

The previous scheme will generate its periodic table updates even if they are not needed at all, i.e. even if no communication is taking place at all. In that case it is not necessary to keep an up-to-date list of current associations, the location of silent hosts is of no interest. Therefore in the following scheme, mobility management is built upon periodic searches, that only occur when communication is intended or is already ongoing. Still the MH is kept free from most responsibility and active participation in order to save the mobile sender’s power resources as well as to solve the problem with outside senders, which is favorable in terms of compatibility to existing network structures. Also it does not involve any new packet types over the wireless link. The functionality is located in the base stations or the portal (Figure 30).

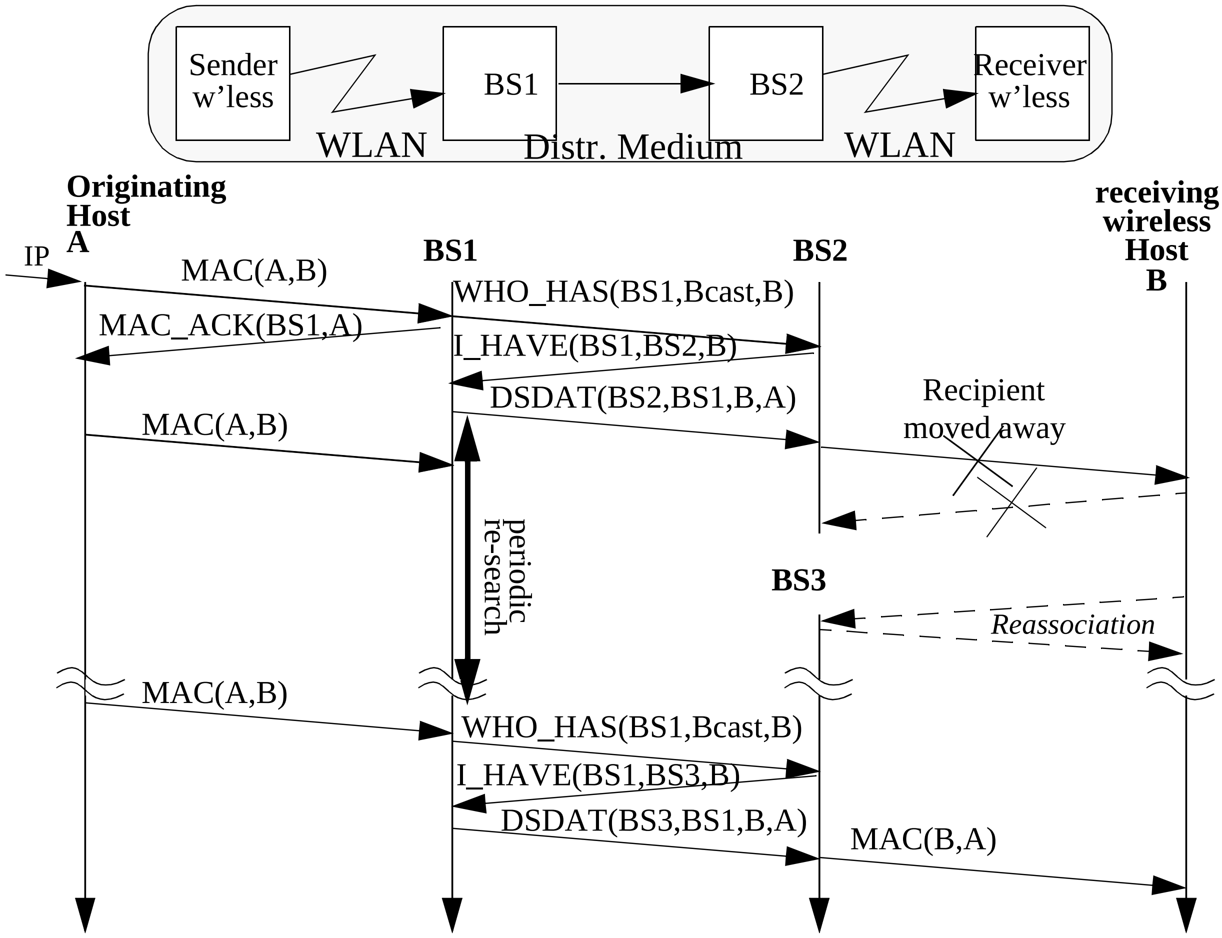


Figure 30:Periodic Search Scheme, Packet Flow

As shown in the packet flow diagram above the sender remains unaware of the actual location of the receiver. The packets do not contain any indication of the location of the final destination. The sender-side base station catches the packet since it knows by looking at its own list of currently served hosts that the receiver is located elsewhere and that the packet cannot reach the receiver within the local cell. The base station therefore buffers the data packet and broadcasts a search to all other base stations. The search is answered as in the previously described scheme by an I_HAVE packet. Upon reception of this reply the buffered data frame is called from the buffer and gets transmitted to the correct base station and subsequently to the receiver. Since this scheme only involves additional control packet transmissions over the distribution medium, which is supposed to have sufficient more bandwidth compared to the wireless link, one may apply an update scheme that builds upon periodic searches. As the receiver moves and reregisters, it cannot be reached until the base station, that is trying to send towards it, initiates one of its periodic searches.

The WHO_HAS and I_HAVE packets now no longer need to be transferred over the wireless link - the respective packet formats are not needed anymore. The search and search reply packets no longer need to carry the information about the host that originally sent out the data packet, since its base station takes care of the search. Therefore the packet formats of both packets can be shortened by one address field:

- Search Request: WHO_HAS(sender, destination, host in question)
- Search Reply on distribution system: I_HAVE(sender, receiver, host in question)

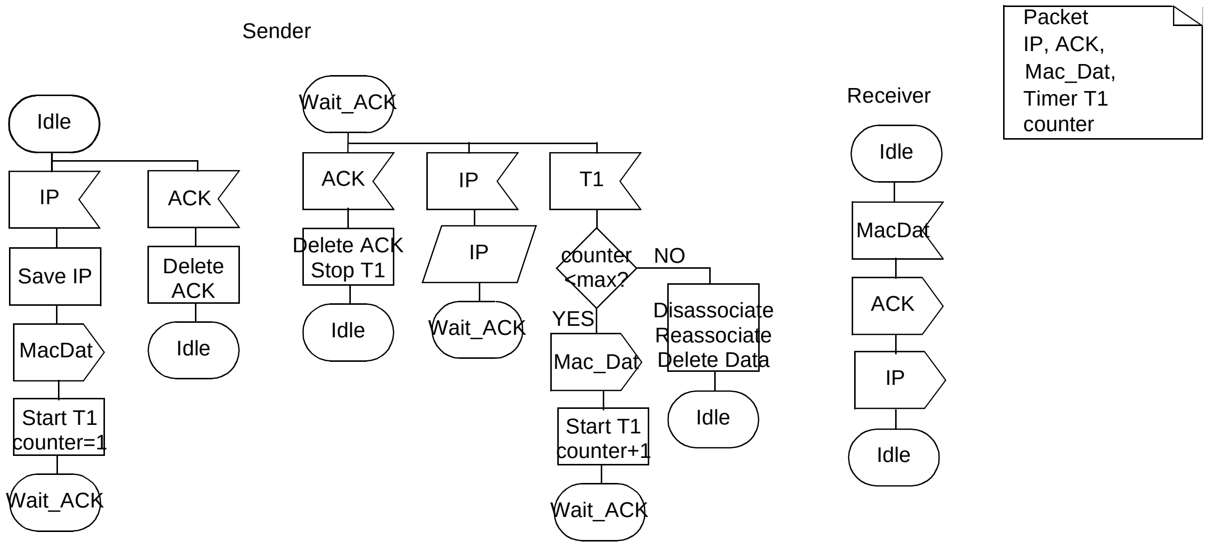


Figure 31:State Diagram Periodic Search Scheme, Mobile Station

The state diagram for the mobile station (Figure 31) remains unchanged from the previous scheme. In the state diagram for the base station (Figure 32) the periodic sending and receiving of tables is replaced by a search process.

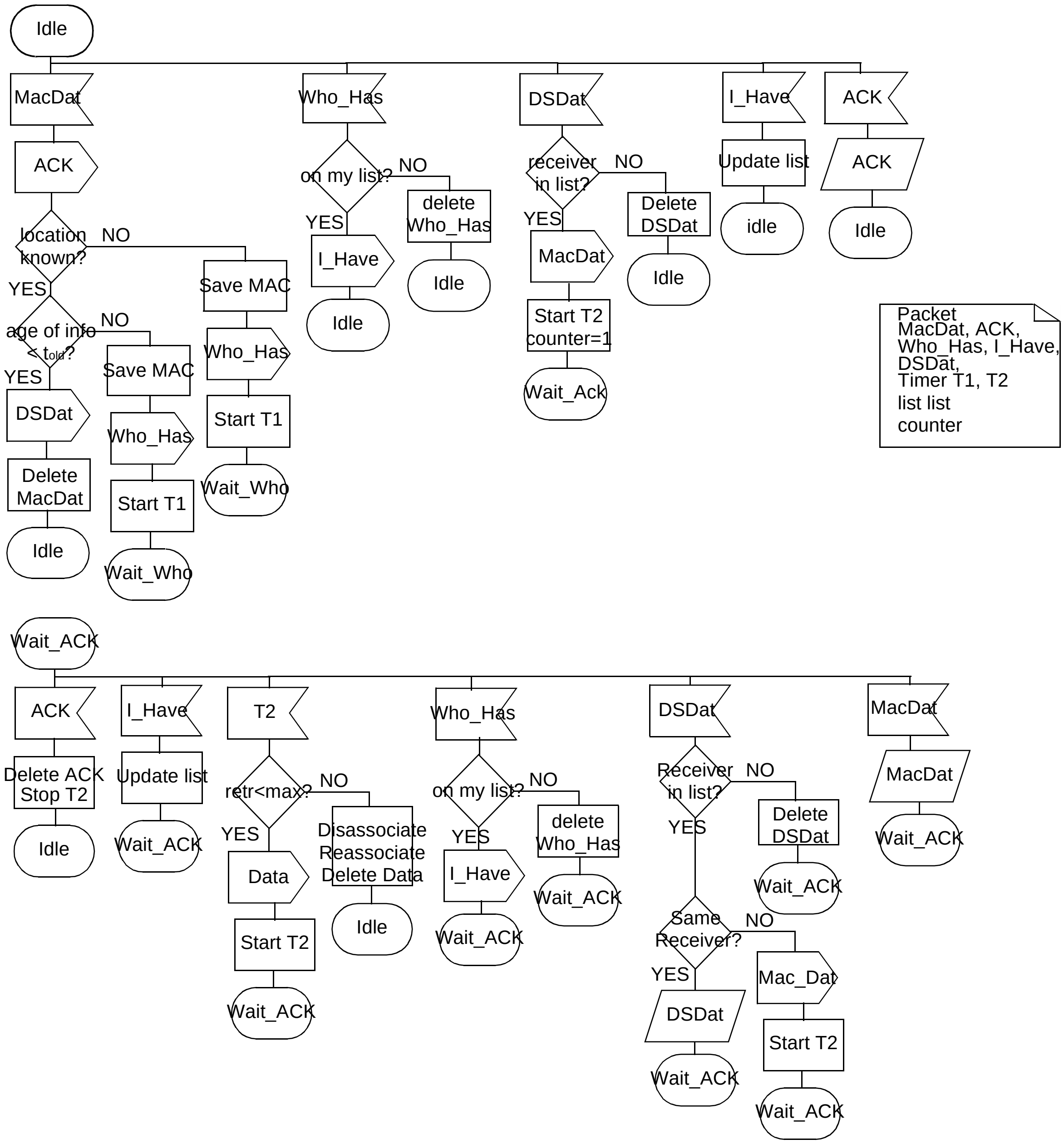


Figure 32:State Diagram Periodic Search Scheme, Base Station, part 1

To be able to integrate the search process into the state diagram a third state Wait_Who has to be introduced, that is entered when a search has been initiated and is exited only when the search has been answered or if it is timed out. MAC packets coming in from the wireless link are now not anymore immediately forwarded or discarded but are stored in a buffer until a search for the receiver has successfully completed or are discarded if the search has repeatedly timed out. This search is periodically repeated by looking at the age of the location information, in order to cope with changing locations of the receiver.

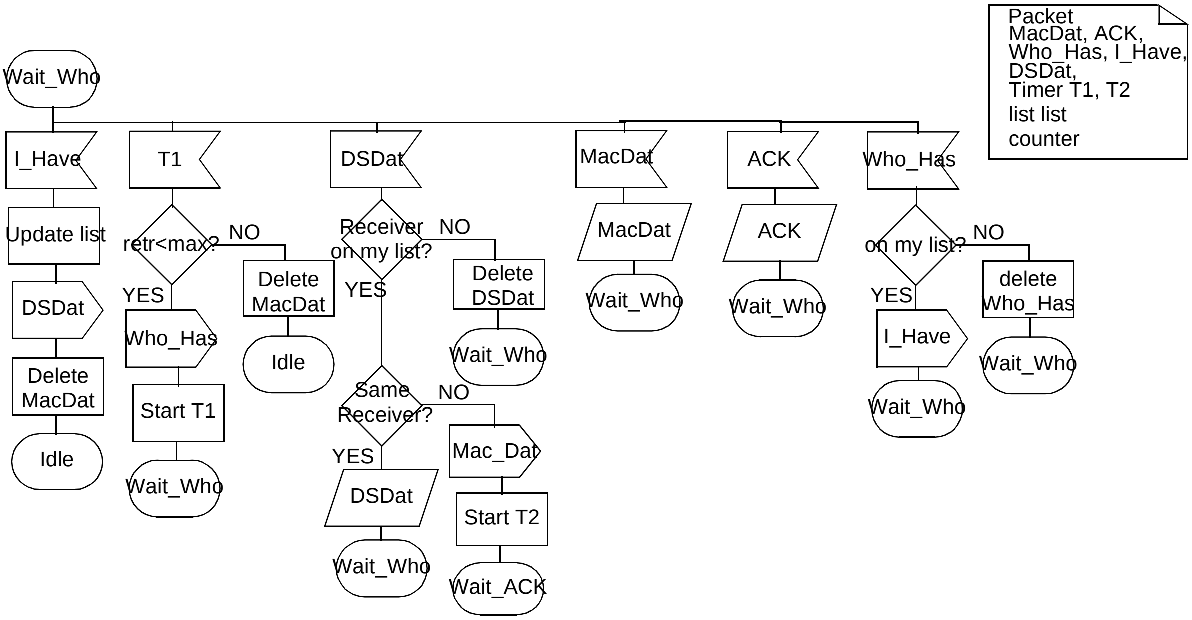


Figure 32: State Diagram Periodic Search Scheme, Base Station, part 2

The use of periodic searches reduces the amount of mobility management information, that has to be exchanged compared to the periodic update scheme by eliminating any mobility management activity when no communication takes place. However as soon as hosts are communicating the search procedures are carried out, independent from the fact if the searched host has moved or not. This problem will be addressed with the next proposed scheme, that applies reactive searches only when a handover occurred, but before this can be attempted the handover event has to be noticed.

4.5. Noticing the Handover Event - Handover Latency

The schemes presented in the previous chapters relied on periodic procedures, that were carried out independently of the actual handover event. Therefore no consideration had to be made on how to detect the handover event. In the following it is attempted to actively detect the handover event and thus allow for reactive procedures. Before such active handover management procedure can be started, some instance of the net has to notice the handover event. This can happen either actively with the help of periodic keep-alive-messages between the base station and the mobile host (“am I still connected” - “yes, you are”) or passively by using timer in either of the two devices, that send an alarm when acknowledgments are missing for a certain time period.

Predictive initiation - beginning the handover processing BEFORE the handover occurs, before the wireless connection to the mobile host breaks - as applied in cellular telephone networks, or even handover “on-the-fly” - i.e. handover management during its occurrence - is not possible in the chosen setup. Technically this requires a certain area of overlapping coverage between the neighboring cells, so that the MH spends a certain time within the overlapping area. This overlapping area may or may not be present in the environment described in chapter 1 and therefore one cannot rely on it. Also signal strength measurement facilities as well as dual-signal reception are needed to determine the stronger one of two received signals, which are features that are not provided in current products. In addition to this the time spent in the possible overlapping area may be very short since the cells only have 10-20 meters diameter. With all these features lacking the MH will detect the loss of connection at earliest AFTER it has lost connection to the old BS and therefore will be unable to initiate any procedure in the old cell.

Active handover detection with the help of periodical messages offers the feature to set the maximal detection time to the desired value (depending on the frequency of the messages) but since it uses additional bandwidth on the wireless link - a lot if fast handover detection is desired - it is not a favorable scheme. The latter scheme applying time-outs performs differently depending on where the first time-out occurs. If one assumes the protocol architecture from the widely used Internet, 3 different acknowledgment packets are possibly available to trigger a time-out if they are missing for too long - MAC layer ACKs between BS and MH (available depending on the selected WMAC protocol), DS ACK between BS and BS (if introduced in the distribution system) and the Transport layer ACKs between MH and Sender (sending from within or outside the WLAN). Figure 33 shows the types of acknowledgment packets and the location where their lack can be noticed. Nothing can be done before the MH has regained connectivity to another BS. The MH **must** take a number of actions on its own without any possible support from other parts of the infrastructure i.e. notice its loss of connectivity, move into the coverage area of a new BS (if the cells are non-overlapping), look for a new BS and synchronize to the new channel in order to be able to exchange possible re-registration parameters. These reassociation procedures are part of e.g. the IEEE 802.11 standard as basic elements in WLAN operation, with or without interconnection to other communication infrastructure. Therefore they need not to be specified in this thesis. At the time of completed reassociation the mobile could actively inform its previous communication partners about its new location. However the remaining hosts in the system may be able to use this period already to detect the handover event, refrain from further transmissions that will be lost anyway and initiate a procedure to regain connectivity with the missing receiver.

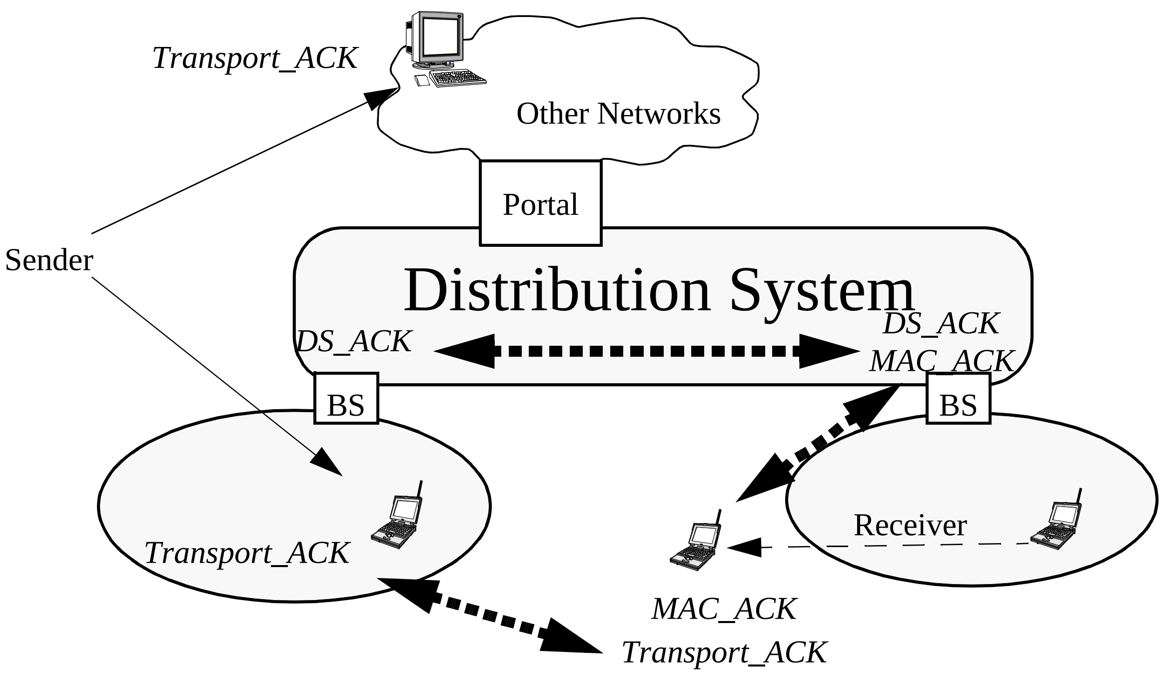


Figure 33:Location of possible Time-Outs due to Lack of Acknowledgments

In the following it is assumed, that both the MAC protocol as well as the transport layer protocol have implemented acknowledgments. For the WMAC protocol immediate ACKs are widely used since they are generally considered necessary to deal with the instability of the wireless transmission medium. If other higher layer protocols (not necessarily limited to transport layer but possibly also connection-oriented network layer protocols) without these acknowledgments are used one may have to replace those end-to-end acknowledgments by cascading hop-by-hop DS acknowledgments upstream. For the WMAC protocol the immediate acknowledgment strategy as known from the IEEE 802.11 standard is assumed, for the transport protocol a TCP-type acknowledgment behavior. In the MH one only has the MAC layer and the transport layer acknowledgments available. The timer constants of those two timers are set to match the needs of the primary purpose of the acknowledgments - error correction - they cannot be optimize freely for the needs of the distribution system in this case of handover detection, without breaking the error correction features. Also the timers are designed to operate in wired environments where one hardly ever sees completely missing packets but rather corrupted packets.The effect for the application is the same -

the packet is dropped in the lower layers - however the reaction on a corrupted packet can be initiated immediately. The need for a reaction on a missing packet - which can justified be interpreted as a sign for a complete station outage - will therefore not occur very often in this environment, which is why the timers can be allowed to operate on a rather long time scale. Therefore one has to look which one may suit best to be ‘hijacked’ for the described purpose.

Usually if MAC Acknowledgments are missing the packet gets retransmitted a certain number of times until a retransmit limit is reached. In IEEE 802.11 this is defined to be 7[1], therefore one would have to wait for 7 successful access attempts to the medium before one can decide that the MH moved out of its cell. Under high load an average access delay in the area of 20 ms has been simulated (for 8 stations in the cell, see [22]), which would result in a handover detection time of up to 7*20ms = 140ms. (HIPERLAN allows 16 retransmissions with an average delay of 60ms for 8 stations under high load) Transport layer timer operate on a much larger time scales - acknowledgments are considered missing usually after about 500 ms (e. g. TCP), then retransmissions are attempted - therefore one would choose the MAC timer for early handover detection.

If a new acknowledgment is introduced on the distribution system its time frame to operate would have to be between the MAC layer acknowledgment timer and the transport layer timer, since letting the timer expire before MAC layer retransmissions may have been done would be counterproductive to the MAC layer retransmissions. Similar initiating a reaction after the transport layer has already attempted retransmissions would not add up to the efficiency of the system. So the trade-off one is facing here is that one wants to notice mobility as soon as possible and as sure as possible, but acting too fast may cause unnecessary actions (retransmissions, reregistrations if time-out is caused by temporary distortion) or harm the functionality on the lower layers.

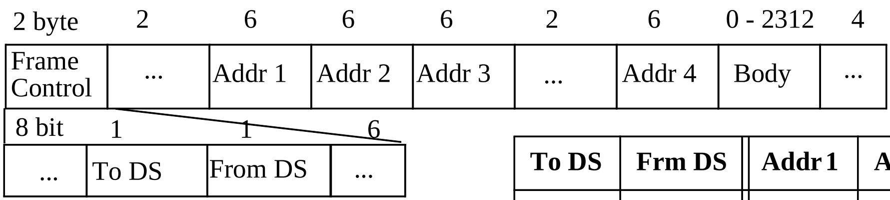
A combined strategy would be to send out a negative acknowledgment over the distribution medium as soon as the MAC layer acknowledgment timer times out. This would use the fastest timer in the system and would allow pointed information of the mobility case to the sending stations and/or its respective base stations. The following section presents the specification of a scheme that applies this mechanism to allow for reactive handover management.

---

1. 7 for short frames, 4 for long frames larger than a certain threshold

4.6. Reactive Search Scheme

The MAC part of the IEEE 802.11 standard has defined its frame format in a way that the station that is initially sending the packet into the WLAN, i.e. the sender or the portal, has to specify a field in the frame whether the frame is destined to a MH in the *same* cell or if it has to go *across* the DS (see Figure 34). To fill in the correct values the originator has to be aware of the location of the receiver, at least to the degree whether it is in the same cell as the sender or if it is in another cell of the WLAN. This information is not supplied by common address resolution mechanisms (e.g. ARP[23]), that do not reflect the location *within* a LAN. An additional mechanism has to provide this information, e.g. an additional search has to be started by the originator before a packet can be issued.



DA - Destination Address
SA - Source Address
BSSID - MAC Address of BS
RA - Immediate (next) Receiver
TA - Immediate (last) Transmitter

To DS	**Frm DS**	**Addr 1**	**Addr 2**	**Addr 3**	**Addr 4**
0	0	DA	SA	BSSID	N/A
0	1	DA	BSSID	SA	N/A
1	0	BSSID	SA	DA	N/A
1	1	RA	TA	DA	SA

Figure 34:IEEE 802.11 Frame Format and Address Field Definition

This location information is needed at the sending mobile host in the WLAN, therefore the search is initiated by this station, as shown in Figure 35.

This is a significant difference to the previous periodic search scheme, where the search process and the accompanying buffering of the frames was done at the base station. This is of course undesirable, since the mobile sending station should not be required to use more of its scarce power than necessary, and since the scarce wireless bandwidth will suffer from the additional control overhead.

Two different events may result in a search, one is carried out at the beginning of a communication, the other one after the detection of a handover. Therefore when a sender first tries to reach a receiver, it sends out a “WHO_HAS” search packet to its base station, that forwards it as a broadcast to all base stations within the E-WLAN. The one base station that is currently serving the receiver replies with a “I_HAVE” packet directly to the respective base station. The reply is forwarded over the wireless link to the sender that issued the search. Now the sender can fill all fields of the frame with the necessary information that describes a full path description towards the destination.

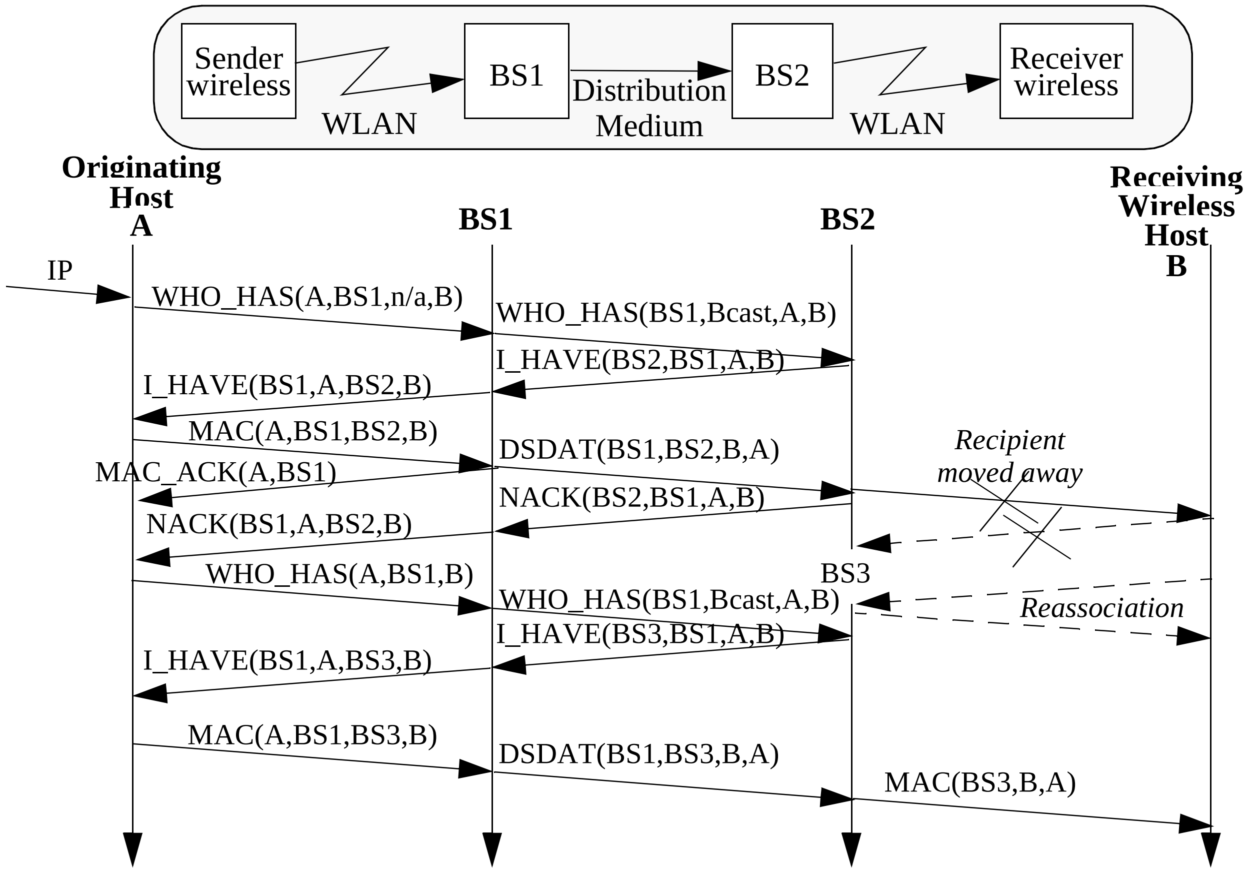


Figure 35:Reactive Search Scheme, Packet Flow

As soon as the sender notices that the receiver can no longer be reached on that path in initiates a new search - eventually repeated searches, if the handover latency period is longer - and resumes transmitting after a successful reply. This however is the crucial issue in this scheme: how does the **sender** notice that the receiver is gone. As discussed in previous Section 4.5. any timer based mechanism to notice loss of connectivity with the receiver, that is located at the sender, is likely to have the largest timescale of all timers in the system. Therefore it is advised to apply a delayed MAC acknowledgment as shown in Figure 22, relying on a transport layer based solution or active notification of the sender upstream as soon as downstream changes become noticed (i.e. negative acknowledgments).

For the reactive search scheme the reactive procedure is based upon negative acknowledgments (NACK) that are issued by the base station, that received a data frame destined to a mobile host in its cell but could not successfully transmit it. The format of the NACK frame is as follows:

- NACK negative acknowledgment: NACK(sender, destination, final destination, lost receiver)

This NACK frame is sent upstream to the base station, where the data frame came from and passed on further up to the sender.

It has to be noted that the address fields in the data carrying packets need to be changed compared to the packets applied in previous two schemes. The address fields in the data packets - both on the wireless part of the path as well as on the distribution medium will have to be readjusted. The new frame formats are:

- MAC data: MAC(sender, destination, next base station, real receiver)

- Distribution system data: DSDAT(sender, receiver, real sender, real receiver)

Additionally the two frames used for the search process are sent both on the wireless link as well as on the distribution system, however with different interpretations of the address fields:

- Search Request: WHO_HAS(sender, destination, searching host (this field is only necessary on distribution system), host in question)

- Search Reply on wireless: I_HAVE(sender, receiver (*i.e. first base station*), base station that has, host in question)

- Search Reply on distribution system: I_HAVE(sender (*i.e. base station that has*), receiver, searching host, host in question)

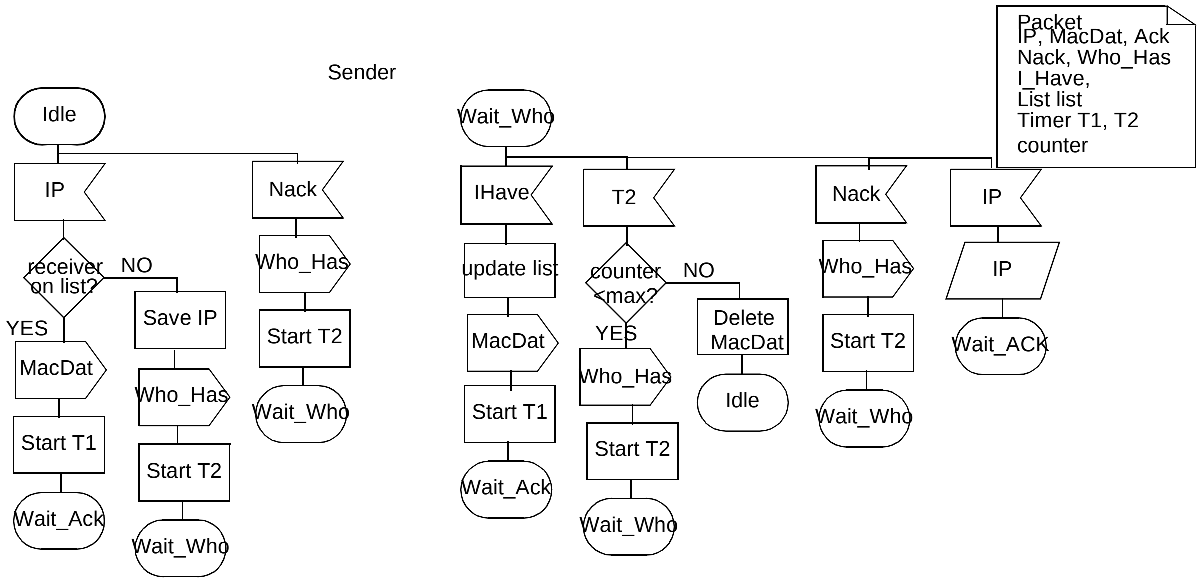


Figure 36:State Diagram Reactive Search Scheme, Mobile Station, part 1

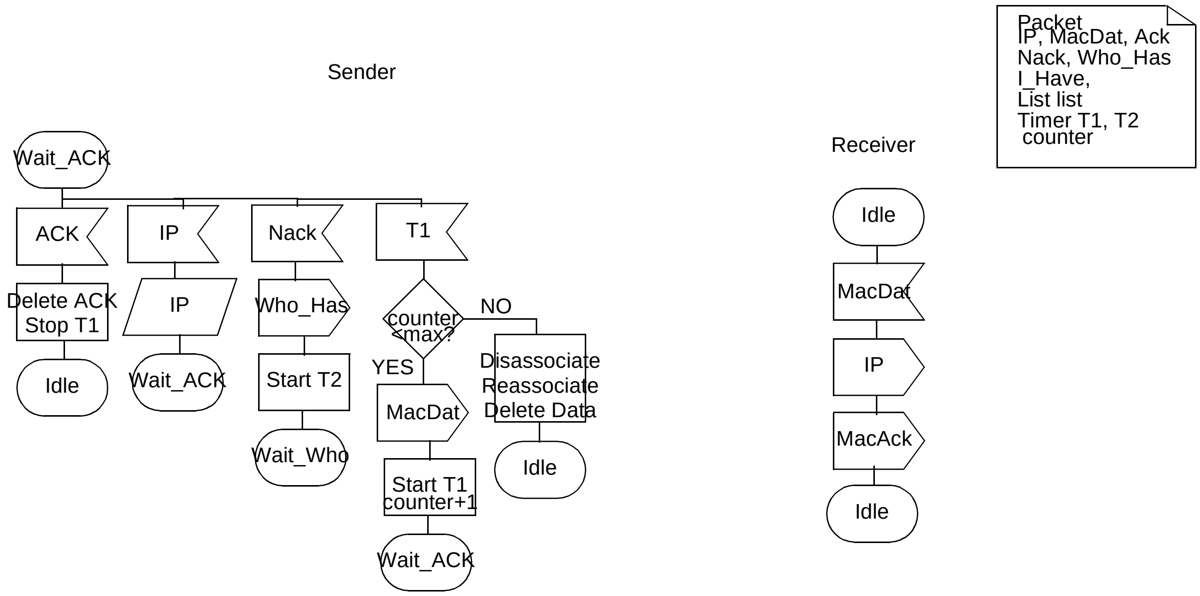


Figure 36: State Diagram Reactive Search Scheme, Mobile Station, part 2

In the state diagram in Figure 36, the state Wait_Who, that was needed in the periodic search scheme in the base station, now has to be moved into the mobile host, since this station now is responsible for the search. Parallel the state diagram for the base station only has two states, *Idle* and *Wait_Ack*. The previously processed search packets as well as the Nack frames, arriving at the sender-side base station are now only rewritten with appropriate address fields and passed on downwards/upwards. Therefore these three frame types now occur in two versions, one for the transmission over the wireless link (indexed with $_{WL}$), one for the transmission over the distribution system (indexed with $_{DS}$).

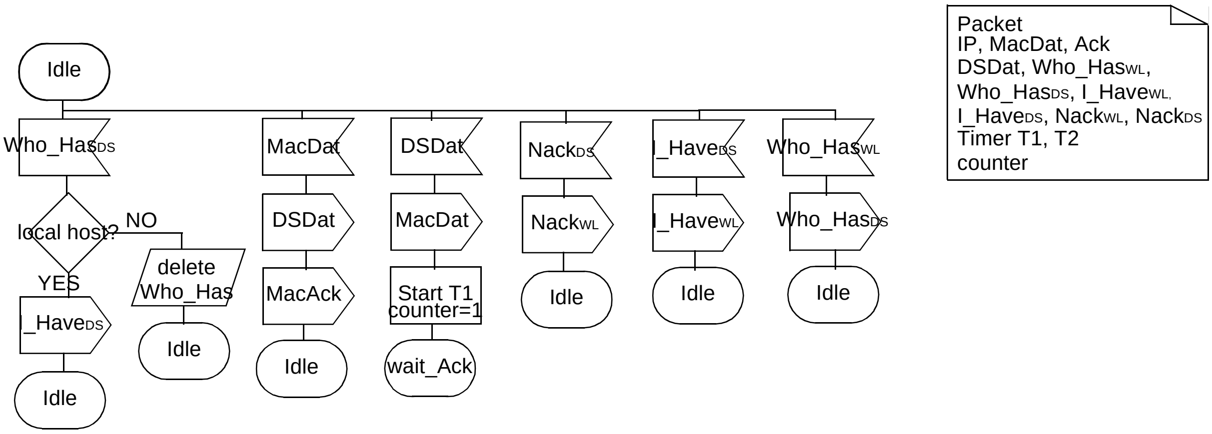


Figure 37:State Diagram Reactive Search Scheme, Base Station, part 1

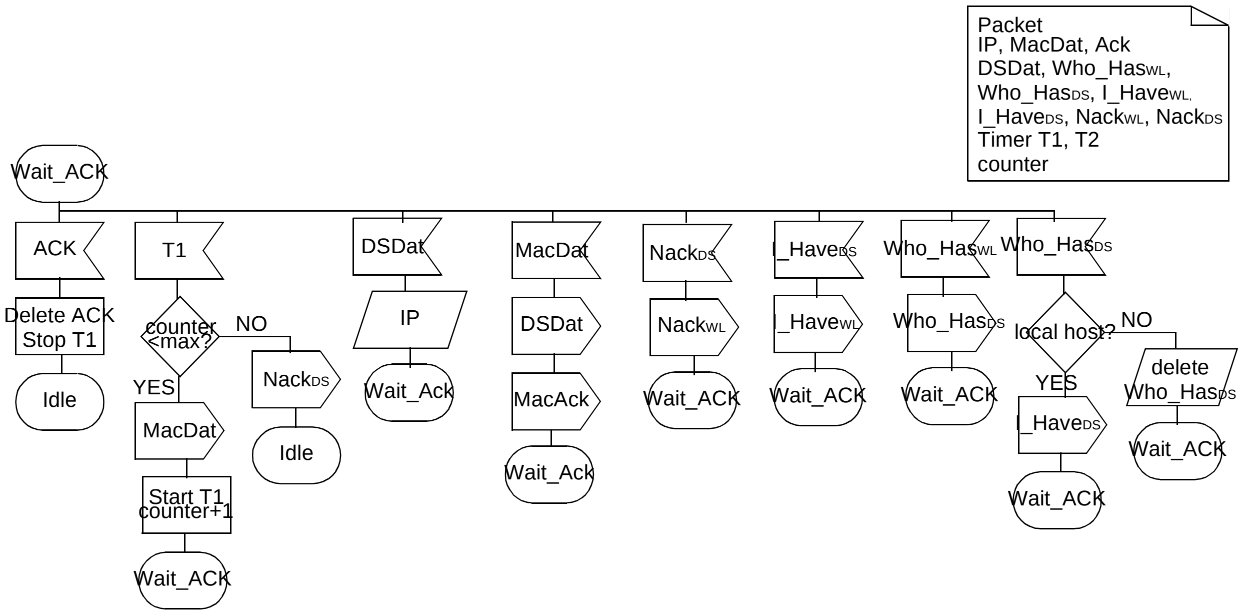


Figure 37: State Diagram Reactive Search Scheme, Base Station, part 2

This scheme also locates a lot of the necessary processing in the sender, that may be a mobile host with limited resources. The additional power consummation will then be counterproductive to power saving features there. Most crucially however the scheme relies on the active participation of the sender. Since hosts sending into the E-WLAN from outside networks cannot be expected to even be aware of the mobility - actually one of the goals of the system elaborated in Section 2 was to provide transparent mobility management to outside hosts. In order to still match this criterion one has to put additional effort in the portal device to act as a proxy for the remote sender, i.e. to *'fake'* the active participation, even though it is not present.

4.7. Multicast based Scheme

The fourth scheme that will be presented in this thesis provides and optimization towards fast handover, in order to minimize the interruption for the reception of a stream of multimedia data. If delay and delay jitter is of larger interest for the desired end-to-end QoS, the interruption caused by a handover should be kept as short as possible.

This can be done with the help of a multicast scheme. The general approach is based on the fact that all packets that are transmitted over the DS, are not only sent to *one* base station but are multicast to a *group* of BS in the neighborhood of the current cell. The mobile receiver, after it moved to another cell, will then find all data, that was sent during the interruption, already waiting for pick-up at the new BS (Figure 38, Figure 39). Due to the complex problem arising with reliable multicast this approach does not provide reliable service - duplicated packets, errors and losses have to be managed on higher layers.

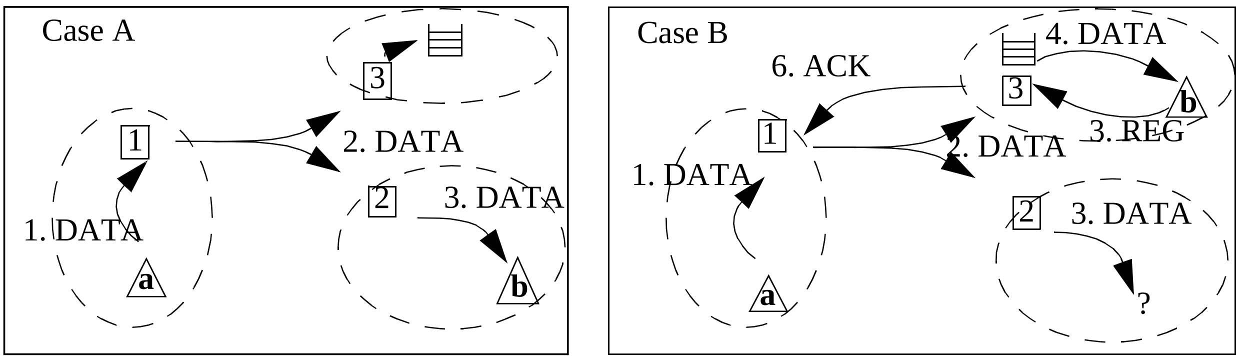


Figure 38:Multicast based Scheme

Figure 38 explains the scheme: in case A, when the receiver **b** is located where it is expected to be, in step 1 the data is sent from sender **a** to BS1. In step 2 the data is sent from BS1 to the multicast group of base stations around the location of receiver **b** (here the group consists of BS2 and BS3). In step 3 the data is sent from BS2 to receiver **b**.

In the case B, when the receiver **b** moves to BS3 the steps 1, 2 and 3 remain identical as in case A, however the transmission in step 3 is not successful. At the same time the reregistration of receiver **b** at BS3 takes place. In step 4 the data that is queued in the multicast buffer at BS3 is sent to receiver **b**. Optional the successful delivery may be acknowledged to BS1. After the handover has taken place, BS1 has to generate a new multicast group around the new location of receiver **b**.

Of course the multicast functionality is not helpful in order to get the location information to the base station, that serves the sender. It only reduces the duration of the interruption. Therefore the multicast scheme is applied as an extension to any of the previously described schemes. In the following the underlying location management scheme is chosen to be the periodic search scheme described in the previous Section 4.4. However the two other schemes may equally be used.

The initial search for the receiver can be seen in Figure 39. In order to inform the base station serving the sender as soon as possible about the location change a negative acknowledgment is sent out from the base station previously serving the receiver. It may then issue a search by itself or wait for the automatically generated I_HAVE packet as shown in the figure below.

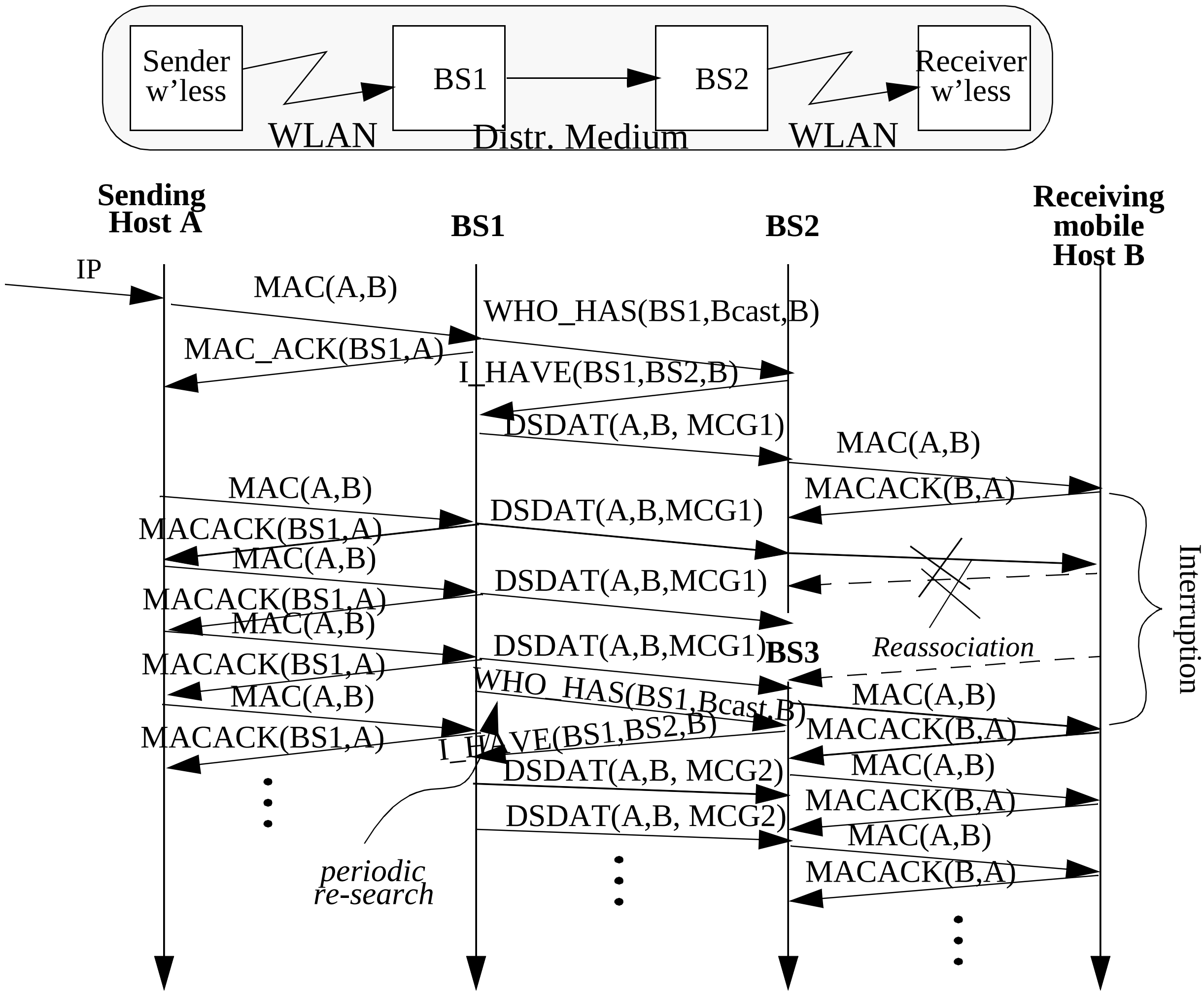
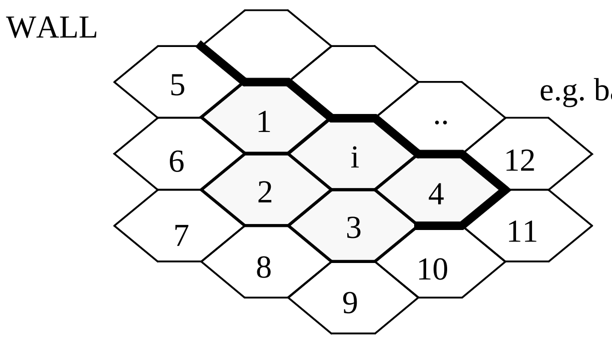


Figure 39:Multicast based Scheme, Packet Flow

This approach requires the appropriate selection of the multicast group around the currently valid base station. In a brute-force approach one may choose to broadcast the packets to all base stations within the E-WLAN. Only the base station in question grabs and forwards the packet to the mobile receiver. This approach has the advantage that the packets do not need to be readdressed on the DS - the broadcast address replaces explicit location addresses. However this approach does not scale well for an increasing number of base stations. If all base stations have to receive all sent packets and store them in their multicast buffer unless they are serving the destination currently, the buffers would easily become very large.

In order to avoid the resource intensive broadcast one will want to choose a subgroup of all base stations by forming multicast groups. The multicast group can be selected depending on the topology of the area - e.g. a host cannot move through walls and can therefore not reach a cell behind a wall - , based on the hosts direction, its speed and its mobility history, The groups can be chosen static or adapted dynamically according to changes in the environment. Figure 40 illustrates possible group lists.



e.g. basestation i member of group i:

g_1={i,1,2,3,4} (all neighbors)
g_2={i,1} (most likely neighbor)
g_3={i,3,4} (neighbors in motion direction)
g_4={i,1,2,3,4,5,..,10}(area coverage)

Figure 40:Multicast Group Selection based on Different Strategies

Many sophisticated strategies in selecting the members of the multicast group are thinkable. E.g. [59] presents a method for selecting multicast subgroups in a similar (connection-oriented) environment. The discussion of this problem is outside of the scope of this thesis, since it would require an evaluation on its own.

Figure 41 illustrates the shortening of the duration of the interruption, that can be achieved with the help of the above described multicast scheme.

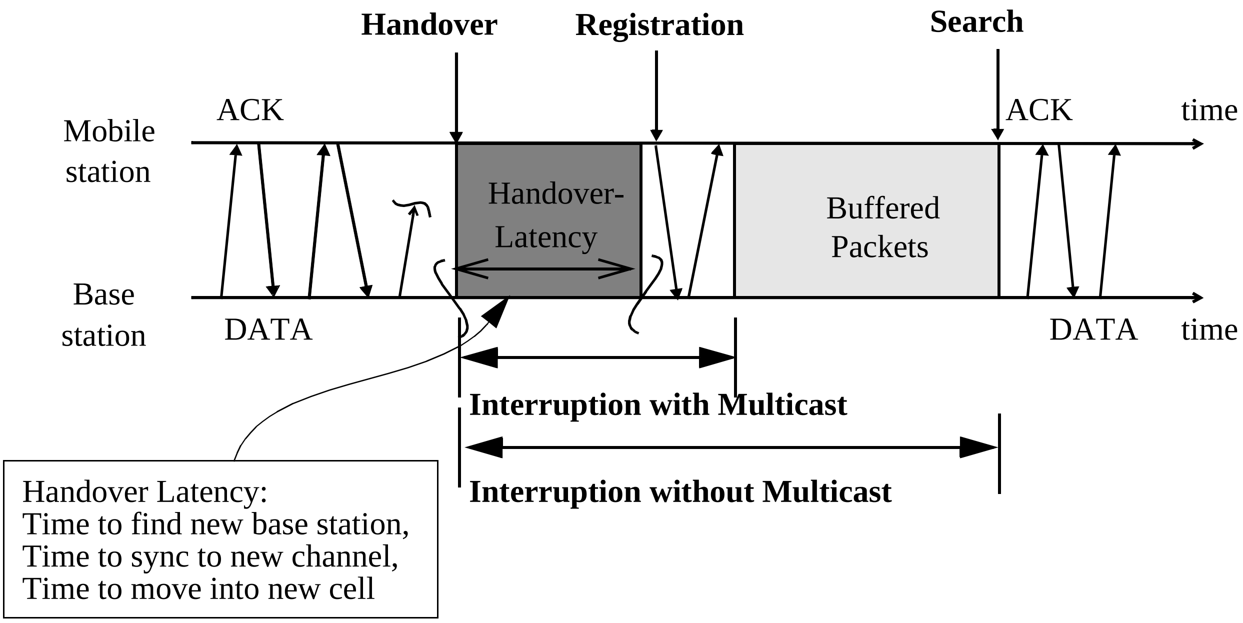


Figure 41:Gain through Multicast Scheme

During an ongoing communication the receiving mobile host looses connection to its base station. After the mobile has moved within reach of a new base station, discovered it and synchronized to the new channel that is used in the new cell it may initiate reregistration. Up to this point the procedure is identical to all schemes, this part of the interruption cannot be shortened. However thanks to the stored frames at the new base station, the mobile receiver will continue to receive data immediately after the reregistration process is completed. Without the multicast scheme it would continue to be unreachable until the next table update cycle is completed (periodic table scheme) or a search has been initiated (periodic search or reactive search).

Figure 42 presents the change in the state diagram for the periodic search scheme when the multicast functionality is added in the base station, where the unaltered parts are marked grey, the new parts are marked black. The changed elements are found in the additional check of incoming DSDat-frames, whether they need to be stored, even if the receiver is not currently served by this base station and in the additional signal after a new association, that results in a check of the buffer for stored packets, that will immediately be transmitted to the newly served receiver. Similarly those changes apply to the other two states Wait_Ack and Wait_Who, that are therefore not shown here.

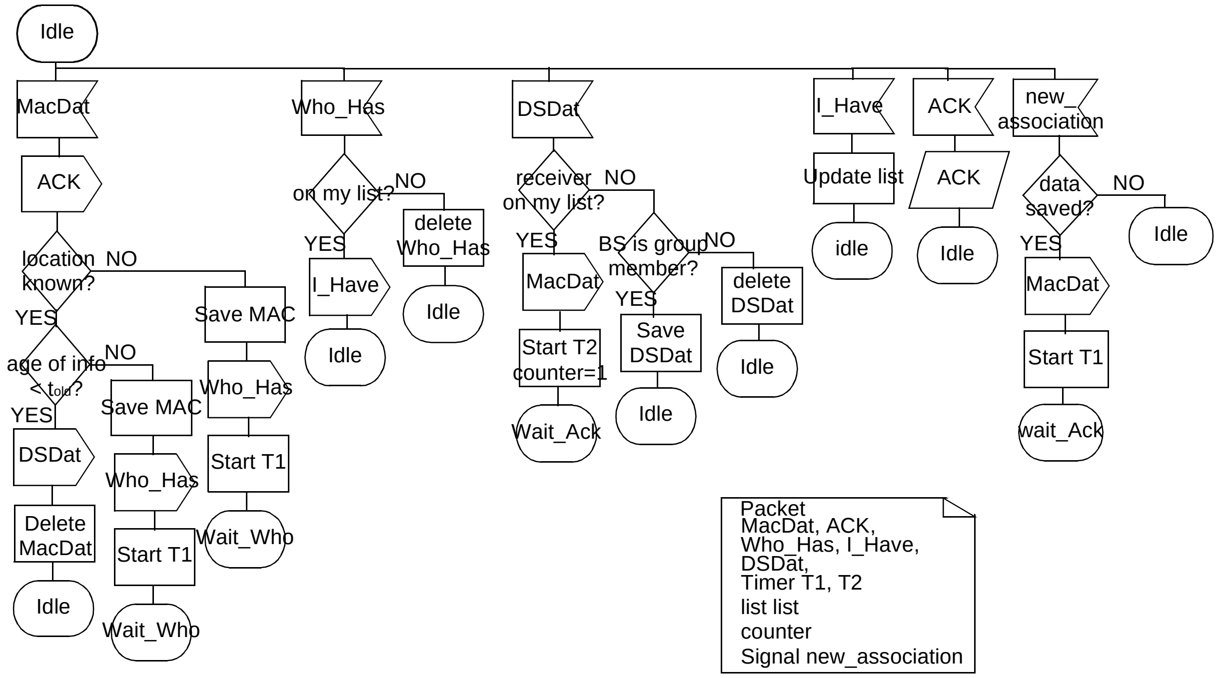


Figure 42:State Diagram Multicast Scheme, Base Station

Chapter 5

Simulation Model

5. Simulation Model

In order to evaluate and compare the schemes presented in the previous chapter a simulation model has been developed. In order to evaluate them one first has to find appropriate criteria for such an evaluation. Common performance parameter like throughput and delay - while giving useful information on the system as well- are not sufficient to give the whole “performance picture”, since those performance parameters that are averaged over a longer time period, do not allow detailed looks on the performance in the particular moments during the handover. Following this discussion on the necessary evaluation procedure, the chapter continues with a description of the applied simulation model. An extensive discussion of the simulation results can be found in the following chapter.

5.1. Criteria for Evaluation

Typically Communication Protocols are evaluated by looking at the achievable throughput over offered load and at the average delay experienced by the communication end points. In the system described in the previous sections however these average figures do not tell the whole story, they do not reflect the behavior during the critical situation of the handover interruption. The following section contains a discussion of an evaluation of the system and develops criteria for this, in order to be able to get comparable results for the performance of the four designs presented in the previous section.

5.1.1. General Evaluation

Generally one may compare the schemes regarding differences in the type of service offered by the system, the robustness of the system and its scalability. The schemes may be compared with respect to the power usage in mobile host and to the number of network nodes involved in rerouting. The qualitative comparison of the different functionalities has mostly been done already in the previous sections where the schemes were presented. The main qualitative distinction is whether periodic schemes are applied (periodic table and periodic search) or if reactive handover management is attempted (reactive search). In the reactive search scheme the sender needs to actively participate, causing additional usage of the scarce resources wireless bandwidth and battery power. The sender here also cannot use packet sniffing on the distribution medium to self-learn other peoples locations. Eventually the portal needs to act as a proxy for outside sender. The multicast scheme is expected to provide the fastest handover, however it imposes higher hardware- and functional requirements (buffers, buffer management, multicast functionality, duplicate detection/handling).

In terms of scalability one has to look at the bandwidth of the distribution medium and at the buffer space. The latter is mainly a problem with the integration of multicast into the system since the packets sent to all base stations in a multicast group have to be stored there. Three influences can be identified that will decrease the availability of those resources for each host in the E-WLAN.

- the number of hosts in the E-WLAN increase
- the available (and subsequently used) bandwidth in each cell increases (by applying a different wireless technology)

- the number of cells increase

If a networking technology is chosen for the distribution medium like switched ethernet that offers inherent scalability feature the issue of scaling the bandwidth may not be an issue at all.

5.1.2.Quantitative Evaluation Criteria

The quantitative evaluation issues concerning the E-WLAN can be grouped under four meta-issues:

- **Impact by providing mobility support features:** delay added to system even if no mobility occurs, additional overhead on wireless link
- **Hardware dimensioning issues:** distribution medium bandwidth requirements, power consumption in mobile host
- **Delay evaluation:** interruption duration, average delay with mobility occurring, number of lost messages
- **System throughput:** end-to-end throughput

Those issues will be described below.

Impact by providing Mobility Support Features

Delay in Static Case

One question arising with the introduction of mobility management into a system is whether the scheme has impact on delay, just by being available, even if it is not used. No obvious reference scenario can be found, to act justified as a fixed reference point for the delay experienced in the different schemes if no handover occur. Therefore only the relative differences between the different schemes in the average delay if no handover events occur can be compared.

Additional Overhead on Wireless Link

In order to simulate the different impact on the wireless links involved the number of control packets that are transmitted over both the sender-side wireless link and the receiver-side wireless link are evaluated.

Hardware Dimensioning Issues

Distribution Medium Bandwidth Requirements

The differences in the requirements for the distribution medium will be evaluated by comparing the number of control packets compared to the transmitted data packets. If the packet sizes are known in the real environment one can use this figure to estimate the traffic load in each environment.

Power Consumption in MH

The simulation model described later in this chapter does not give any information on the power usage in the mobile host. Too few is known about the power usage of the different procedures running in the network interfaces in the mobile host. E.g. there are differing opinions whether the most power-intensive task in such a wireless interface is transmitting data or keeping stand-by mode up. Furthermore even the process of transmitting data consists of separate subtasks, where the task of tuning the equalizer has been found to be significantly more power-consuming than the actual transmission of the data itself, which is why different packet sizes can be neglected for power consumption evaluation and be replaced by plain packet counting.

Delay evaluation

Interruption Duration

In order to compare the behavior in the critical moment of the handover the mean interruption duration is evaluated. This includes the time to notice mobility case, the time to initiate reaction, the time to get updated info to sender, and the time until the first packet is sent.

Average Delay

Generally of interest in the evaluation of network architectures is the average delay that is experienced by the data packets sent from sender to receiver.

Number of lost Packets

The rate of lost packets during a handover is simulated and put into relation to the total number of transmitted packets. This figure of course is highly dependent of the duration of the handover interruption but - as discussed above - it may well remain unchanged of it if the source is informed about the event and is therefore blocked from sending.

System Throughput

System Throughput

The overall system throughput allows to compare the influences of the different schemes on the data stream.

5.2. Simulation of Distribution Systems

5.2.1. Scenario

The model consists of 10 cells represented by the 10 base stations. The base stations are interconnected over a bus-type distribution medium. The simulation setup is shown in Figure 43 with one mobile station β sending stationary to the other mobile station α that is moving around freely through the cells.

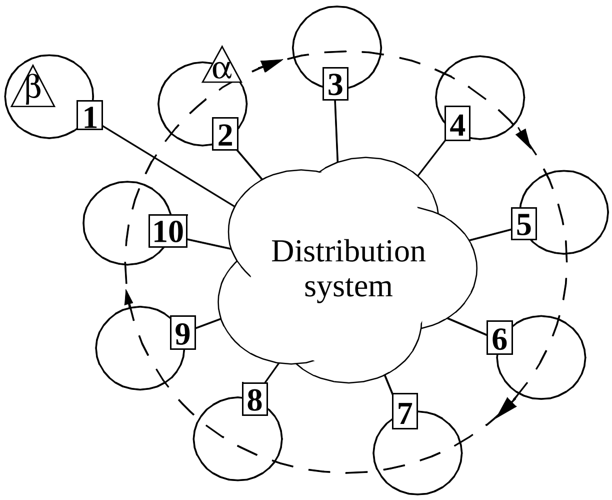


Figure 43:Simulation Scenario - one Mobile Receiver, one Static Sender, 10 Cells

5.2.2. Simulation Tool

To carry out the above described performance evaluation, a simulation model of an E-WLAN system of the described type has been developed. The simulation package Ptolemy version 6.0 [1] has been used for this. Ptolemy is a simulation environment developed at the University of California in Berkeley[51]. It has a strict modular and hierarchical structure and is very flexible since it offers several different simulation domains. Among those are a flow oriented domain for signal processing, a discrete event domain and others. For the simulation the discrete event (DE) domain has been selected since it implies the connection of events and actions from its SDL-

1. Version 7.0 has been released recently but was not usable in the available computing environment at the time of the simulations.

approach (states and state transitions through events). States and state transitions have to be created separately. There is one fundamental difference however: SDL initiates from a state and looks at different events, the DE domain implements this assignment the opposite way - in the case of an event an action and a state transition is initiated depending on the state (Figure 44)

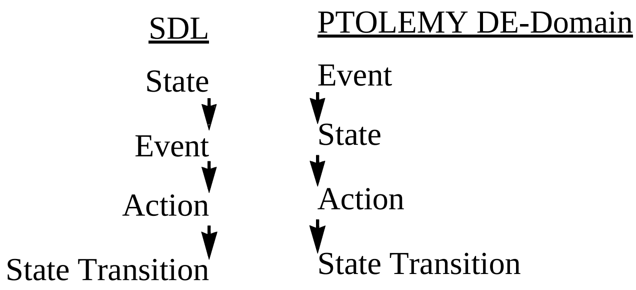


Figure 44:Mapping of SDL to the PTOLEMY DE-Domain

A Ptolemy model that can run a simulation is called a universe. A universe can consist of stars and/or galaxies. The latter may again either consist of stars or of galaxies. At the lowest hierarchical layer of the model only stars can be found.

5.2.3.Simulation Models

In this chapter the different models developed to simulate the E-WLAN system are presented. The most important and active nodes in the E-WLAN model are of course the mobile host and the base station. The two most essential elements in those two hosts are the protocol machine for the distribution system, that are of course different in the two hosts. Together with simple models for the media access and the physical layer of a wireless connection a model for the E-WLAN can be built up as shown schematically in Figure 43. Figure 45 shows such a full E-WLAN model.

On the left side of the figure the symbols for the two mobile hosts can be seen that receive and send packets only to/from the base stations. The 10 symbols for the base stations, representing the cells, can be seen in the center of the figure. To the right side of the base stations the distribution medium is attached, as well as one of the statistics stars, that counts data and control packets transmitted over the distribution medium.

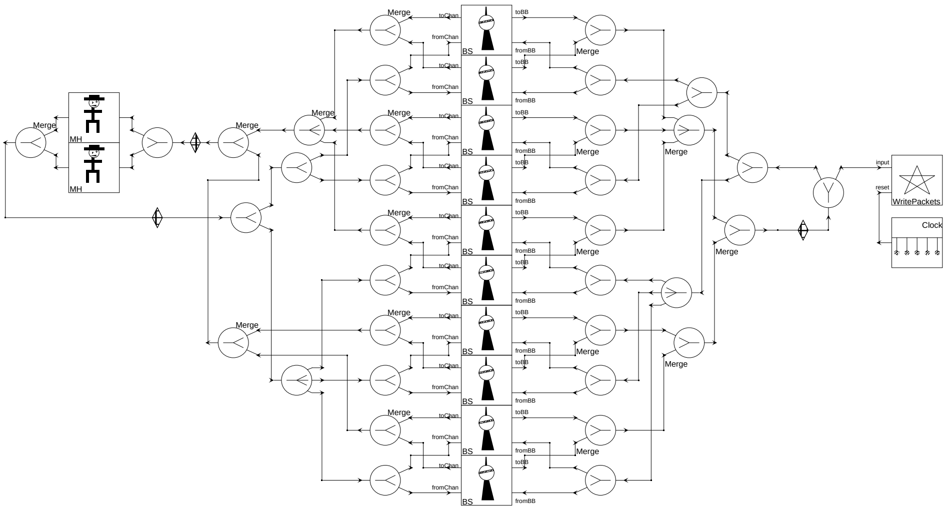


Figure 45:Overall E-WLAN Model with 10 Cells and 2 Mobile Hosts

According to the Ptolemy hierarchy the two mobile hosts are each represented by a galaxy that itself consists of several galaxies (consisting of stars) that represent each of the layers of the OSI-layered protocol stack (Figure 46)- if present at all, i.e.

- IP - the application galaxy, serving for load generation and as a packet sink for end-to-end evaluation (Figure 49)

- DS - the distribution system galaxy, providing the mobile hosts distribution functionality, different for each scheme (Figure 53, Figure 57, Figure 55, Figure 60) The schemes are applied by simply replacing the current galaxy with the galaxy to be applied.

- MAC - the MAC layer providing medium access functionality (Figure 51)

- PHY - the wireless physical layer (Figure 52)

The mobile host is completed by a handover-event-generator that generates handover events at a given mean rate with exponential distribution around that mean value. Furthermore incoming and outgoing ports towards the wireless channels are provided.

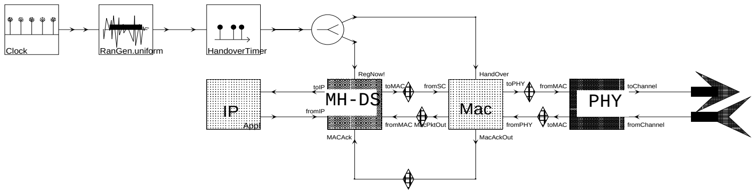


Figure 46:Model of Mobile Host

The base station similarly consists of several galaxies (Figure 47):

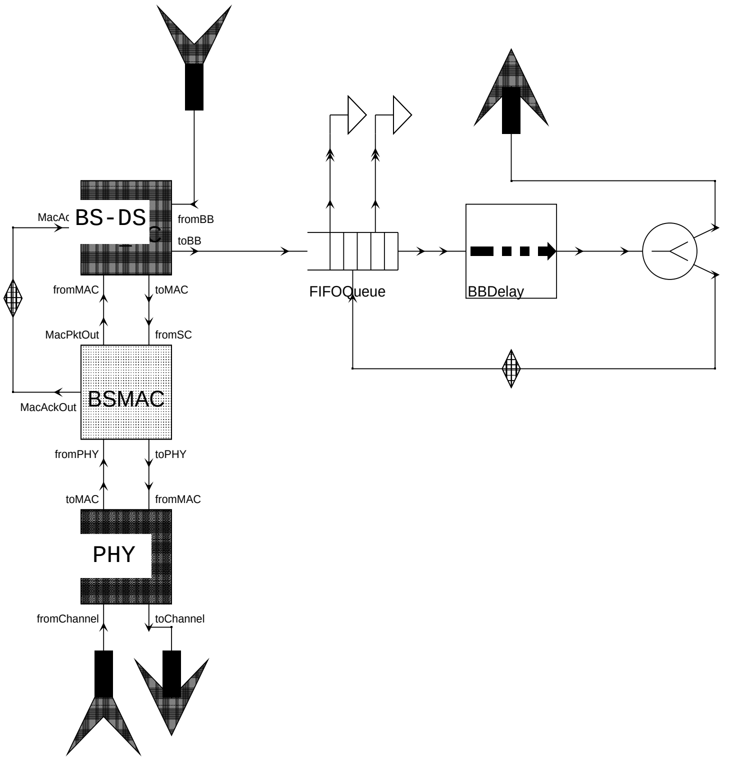


Figure 47:Model of Base Station

- DS - the distribution system galaxy, providing the base stations distribution functionality, different for each scheme. (Figure 54, Figure 58, Figure 56, Figure 59) Similar to the mobile host the different modules for the distribution system simply need to be replaced to use a different scheme.

- BSMAC - the MAC layer galaxy for the base station (Figure 50)

- PHY - the wireless physical layer (Figure 52), identical to the PHY galaxy of the mobile host

The wireless path in the model of the base station is completed by the two ports for incoming and outgoing transfers.

The layered approach for the communication over the wireless link can also be seen in Figure 48, where the layering of the primitives exchanged over the wireless link is shown The protocol stack for transmissions over the distribution medium has been kept very simple. Packets are delayed according to their length and to the bitrate on the distribution medium, without having to go through a MAC galaxy or a PHY galaxy. This implies that each base station has assigned the same full bandwidth available for its purposes. If the actual system applies e.g. switched ethernet this assumption is justified. If this condition is not the fact, however if the bandwidth of the distribution medium is significantly higher than the bandwidth in the wireless islands and therefore the load condition in the distribution medium is generally low, the limitation will not have much impact on the bandwidth available for each cell.

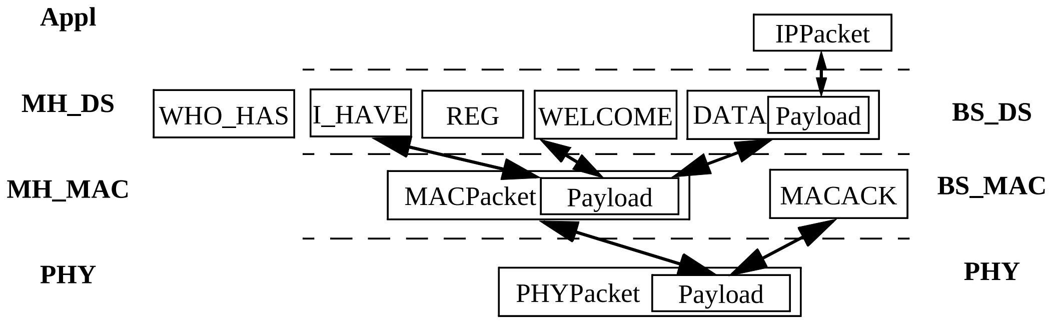


Figure 48:Exchange of Primitives between the Galaxies

Another important fact in the model is that even if the sender and the receiver are in the same cell they can neither communicate directly nor directly over the base station without going over the distribution system as it is optionally selectable in the two WLAN standards ([3],[4]) and widely applied in most WATM schemes. As a simulation detail in this context it has to be added that the base station always assumes that the receiver is located elsewhere and initiates e.g. a search, that it will eventually answer itself.

The application galaxy (Figure 49), operating as the load generation module of the system, basically consists of a star that receives at Poisson distributed times an increasing sequence number. It then sends packets with these sequence numbers to the distribution galaxy. It also collects the incoming packets on the receiving end of the communication and feeds them into two statistics stars.

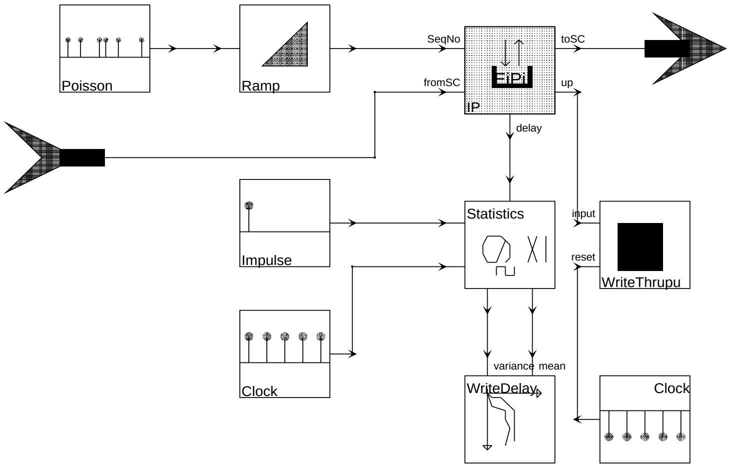


Figure 49:Model of the Application Layer Galaxy

The MAC layer galaxies for the mobile host (Figure 51) and for the base station (Figure 50) differ to a certain degree, mainly to the point that the MAC galaxy of the base station does not have to keep track of the current location. In the path going from the distribution system down to the physical layer both encapsulate the DS packet into a MAC frame and queue it into a FIFO queue and subsequently in a delay star. The delay star adds a fixed amount of delay to each frame to implement the access delay on the MAC layer. The value chosen for this delay is taken from simulations of the IEEE 802.11 medium access protocol [22] and reflects the average delay experienced for an average configuration (8 mobiles in one cell) under high load. All packets arriving from the physical galaxy are passed into a MAC router star that filters out all data packets destined for this particular host. All MAC acknowledgment packets destined for this host that are arriving in response to a previously sent data packet are directly passed on to the MACACK port of the DS galaxy. The data packets are taken over from the MAC_Machine star, that performs differently for the base station (Figure 50) and the mobile host (Figure 51). In both cases the MAC_Machine generates acknowledgments for each packet and feeds them back into the outgoing queue in the reverse direction. The MAC_Machine in the mobile host galaxy (Figure 51) also receives the handover signal from the handover event generator. With the help of this signal the MAC galaxy stays aware of the current location and can reject packets arriving from the no longer valid, old base station. In order to implement a certain handover latency period, that occurs after every handover due to

- spacial distance between cells
- time to synchronize to new channel

- time to find new base station

a timer is started after each handover that blocks the MAC_Machine from receiving or generating **any** message during this period.

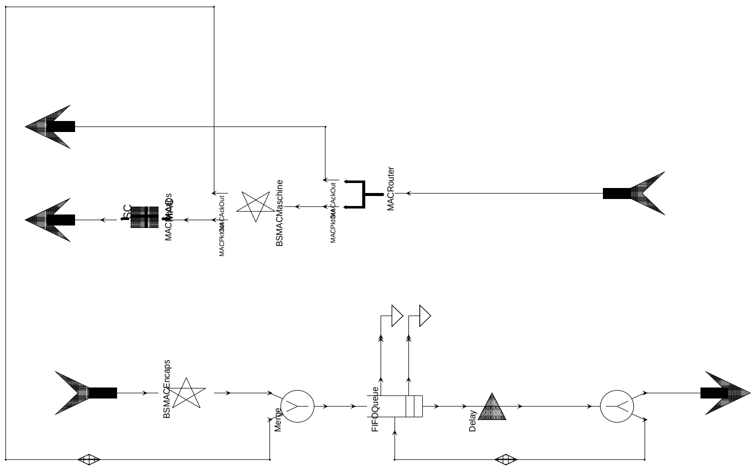


Figure 50:Model for wireless MAC Layer in Base Station

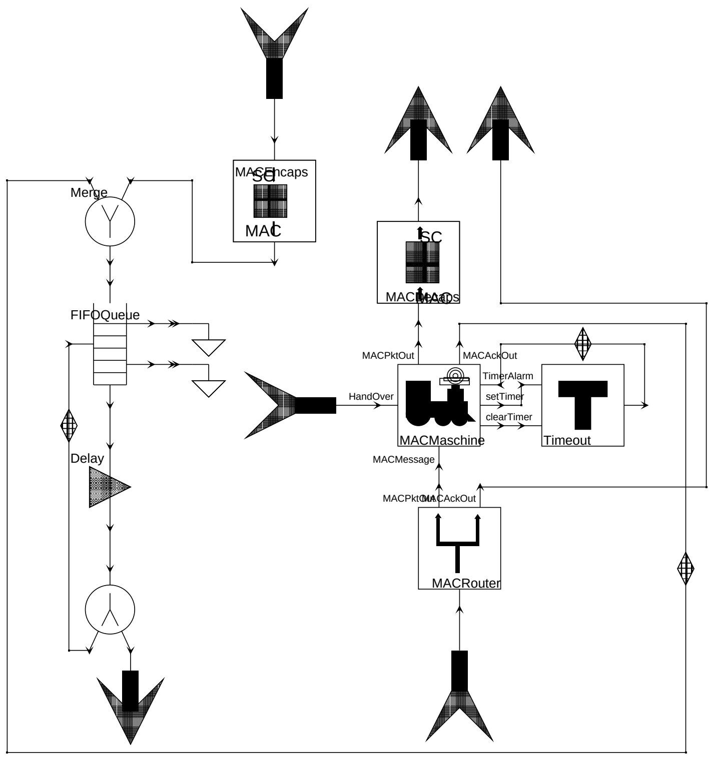


Figure 51:Model for wireless MAC Layer in Mobile Host

The galaxy for the physical layer of the wireless connection (Figure 52) again is kept rather simple. The stars ‘PHYEncaps’ and ‘PHYDecaps’ encapsulate or decapsulate the packets from the

MAC layer or from the PHY layer respectively. On the path from the MAC galaxy to the wireless channel the MAC frame is queued in a FIFO queue and passed through a delay star that adds delay according to the bandwidth of the wireless channel. The FIFO queue takes care that at each moment only one packet is transmitted over the wireless link. Uplink and downlink are separated. The downlink as the connection from the base station to the mobiles is only accessed by the base station, the uplink is shared between the stations in the cell.

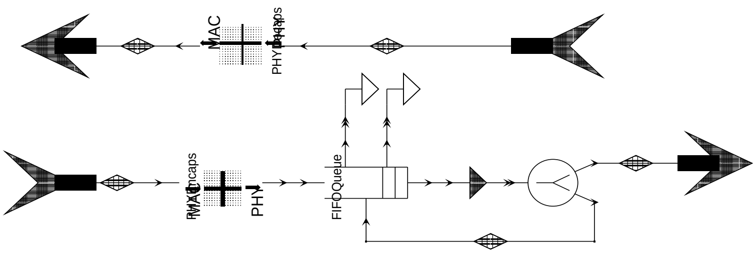


Figure 52:Model for wireless Physical Layer in Mobile Host and Base Station

The galaxies described so far are identically applied in all the four approaches that shall be evaluated in this model. The differences in the functionality can all be found in the distribution system galaxies, that are developed separately for each approach in both the mobile host and the base station. In order to simulate one of the design variants one simply has to replace the galaxy for one design in both participating galaxies with another type galaxy. This approach ensures a simulation environment that provides comparable results from the different simulation runs.

In the following the stars that are used to build up the distribution system galaxies for both the base station and the mobile host will be described. The description will start with a full description of the **periodic table scheme** (see Section 4.3.). Followed by that will be a description of the other three schemes by working out the differences the have to be made in order to simulate the different functionalities.

As can be seen in Figure 46 the DS galaxy in the mobile host receives application messages from the application galaxy, MAC messages from the MAC galaxy as well as the handover event signal. It sends messages out in the reverse direction to the application galaxy and to the MAC galaxy. In the DS galaxy for the periodic table scheme in the mobile host (Figure 53) the handover event signal is passed immediately to the star that receives incoming frames from the lower layers in order to block it from receiving any further messages. The signal also triggers the manager star to signal to the registration star that it shall start a re-registration process with the new base station.

Upon reception of a welcome-packet arriving from the base station, the registration star signals the completion of the registration back to the manager star that in turn informs the transmit-star and the receive-star of the new association and that new packets may now be received and transmitted from or to the new base station. The re-registration process is secured by a timer that causes a retransmission of the registration request if no welcome packet has arrived during the lifetime of the timer. The messages coming in from the application galaxy are passed into the transmit-star that forwards them towards the base station. It also keeps track of its previous transmissions with the help of the MAC layer acknowledgments and a MAC acknowledgment timer. In the case of a successful transmission the transmit star signals the manager its readiness to process the next packet, that in turn gets called from the FIFO queue. The remaining stars in the MH periodic table DS galaxy serve as a router for packet type filtering and as a packet counter for traffic statistics on the wireless link, both upstream and downstream.

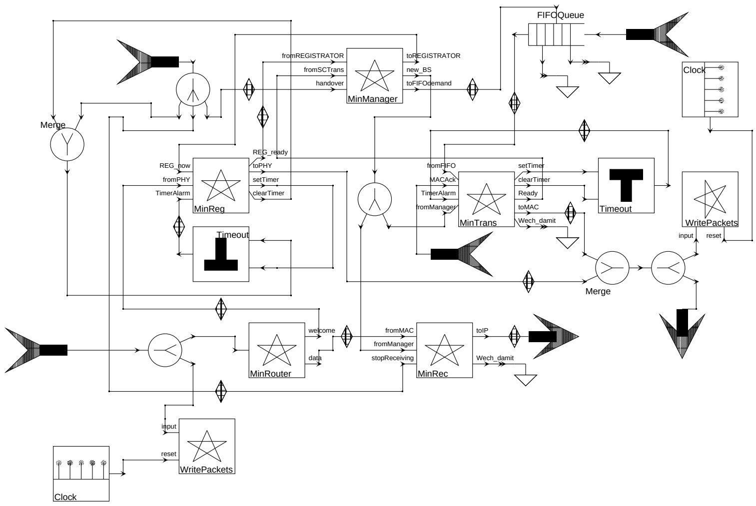


Figure 53:Model for Distribution Layer in Mobile Host, Periodic Table Scheme

The corresponding DS galaxy in the base station (Figure 54) has incoming and outgoing ports to both the wireless link and the wired link to the distribution medium. As in the mobile host the MAC layer acknowledgments, that arrive from the mobile host in response to a previous transmission are passed directly into the DS galaxy. As in the mobile host the acknowledgments are used to keep track of successful transmissions and to trigger the demand of a new packet that is queued for transmission. All other incoming messages are first filtered by type in the Front-router (from the wireless link) and the Back-router (from the distribution medium). The core stars of the DS galaxy in the base station are the home location register (HLR) and the foreign location register (FLR)[2]. The HLR maintains a list of the mobile hosts that are currently served by the base station.

Therefore it has to receive the registration requests and has to generate replies on them. It removes entries from its list upon reception of a negative acknowledgment indicating a failed transmission that has been passed on from the MAC_ACK star.The FLR maintains a list that contains entries that map a mobile host to its current base station. In order to enable the FLR to gain this information in the periodic table scheme the HLR stars in each base station periodically broadcast their table of currently served mobiles to all other base stations. Now if data arrives from a mobile host either the FLR has no current information about the location of the destination in its table available and drops the data, or it has an entry in its table and forwards the message to the appropriate base station over the distribution medium. At the receiving base station in the HLR star the packet again is either dropped, if the mobile has moved away in the meantime and the entry has subsequently been removed, or its is sent out further over the wireless link to the mobile receiver.

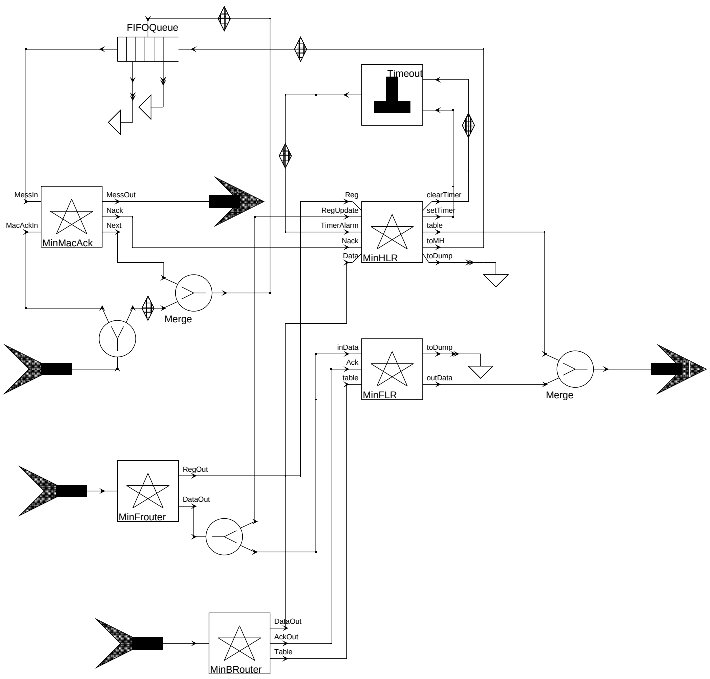


Figure 54:Model for Distribution Layer in Base Station, Periodic Table Scheme

For the **periodic search scheme** (described in Section 4.4.) the DS galaxy in the mobile host remains unchanged (Figure 55). The changes are only incorporated in the base station (Figure 56).

2. Even though the terminology here is taken from GSM-style wireless telephone network architecture the functionality is different.

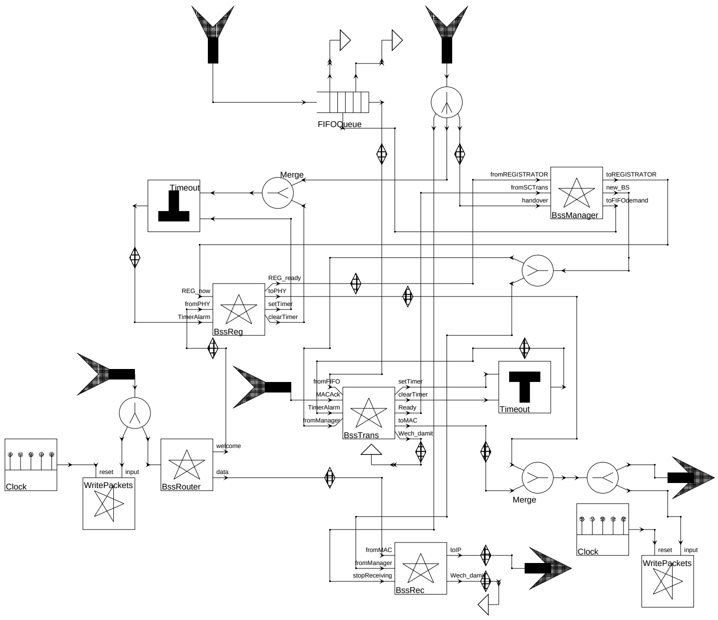


Figure 55:Model for Distribution Layer in Mobile Host, Periodic Search Scheme

The mechanisms to distribute the connectivity tables over the distribution medium is removed in the HLR and FLR star. In order to get the information, which HLR is currently serving the destination host a search scheme is added. The FLR star, upon reception of a packet, destined to a host for which it does not have up-to-date information already available, stores the data packet in a buffer and sends out a broadcast “WHO_HAS”-packet to all base stations.

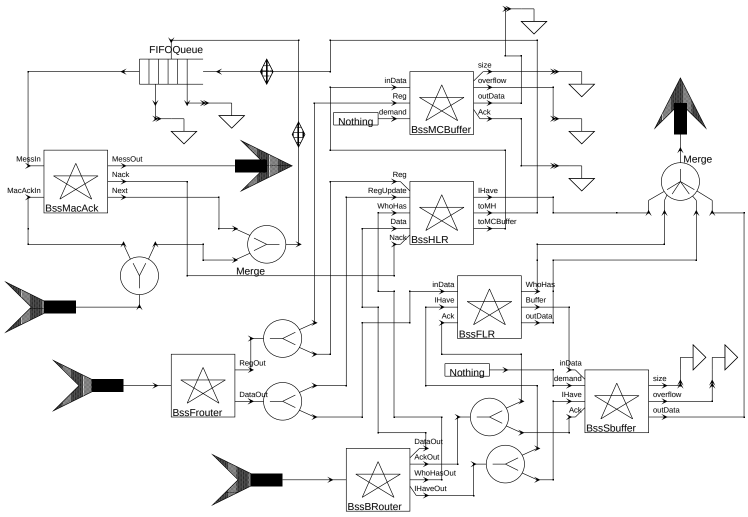


Figure 56:Model for Distribution Layer in Base Station, Periodic Search Scheme

The current base station, that has the destination in its HLR table replies with a "I_HAVE"-packet. The packet is then called from the buffer and sent directly to the base station in question, where it is received by the HLR star and forwarded onto the wireless link. Since this scheme does not incorporate any acknowledgments over the distribution medium (see Figure 22) failure to transmit the packet over the wireless link is not signalled to the FLR module of the base station. Therefore periodic searches have to be initiated in order to be able to take invalid entries from the FLR list.

As discussed before in Section 4.6. the search functionality in the **reactive search concept** (described in Section 4.6.) is preferably moved to the mobile station, as shown in Figure 57. The model for the reactive search scheme therefore has to be changed in some points in order to integrate the search functionality in the mobile host instead of in the base station and to add the negative acknowledgment needed to notify the mobile of the lost association on the receiving end (Figure 57).

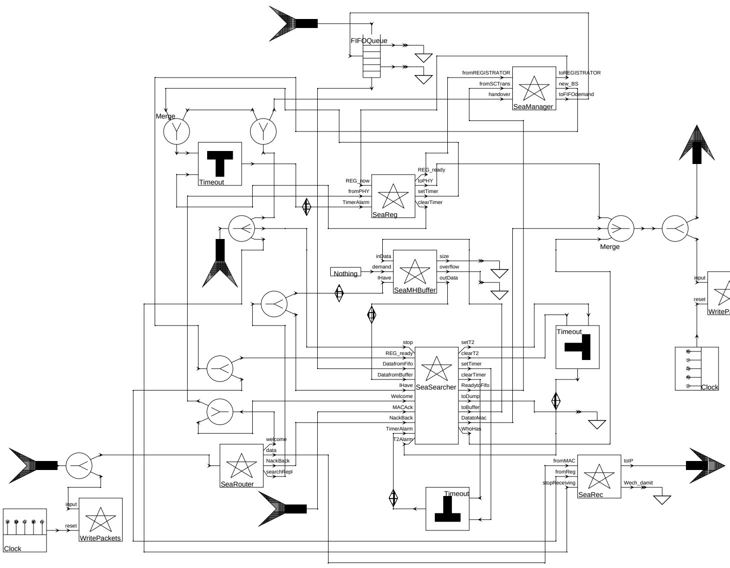


Figure 57:Model for Distribution Layer in Mobile Host, Reactive Search Scheme

This is required so that the sender can determine the location of its destination itself. Compared to the model presented in the description for the periodic search scheme, the send-buffer star has to be moved from the base station into the mobile host, as well as, of course, most of the functionality for search initiation from the FLR star. The searcher-star is added into the DS galaxy. The searcher star now has to issue WHO_HAS packets upon reception of a packet destined to an unknown host, while storing the packet in its attached send-buffer. The search packets are sent over the wireless link to the local base station where they are broadcast over the distribution medium and answered as before. The model for the base station (Figure 58) is reduced by much of its functionality - the remaining tasks are mainly readdressing and forwarding packets as well as broadcasting and answering search requests. The FLR module passes the replies from the distribution medium on to the mobile host. As before the searcher star updates its table with the newly gained information, calls the packet from the buffer and transmits it with the full path written in its address fields.

In order to allow reactive (and thus only the necessary) searches over the wireless link an negative acknowledgment functionality has been added: unless notified otherwise the sending mobile host considers its destination to be still located where previously found. If a transmission to the mobile receiver fails, this is noticed by the old base station previously serving the destination. It then generates a negative acknowledgment, that is sent to the base station of the sender and from there passed on to the mobile sender. At this time a new search is justified and therefore initiated. This is shown in Figure 58 where the NACK packets are generated, which is transmitted upstream to the sending base station. This base station passes the NACK packet over the wireless link to the mobile sender, where it is received and processed in the FLR module, that initiates a new search.

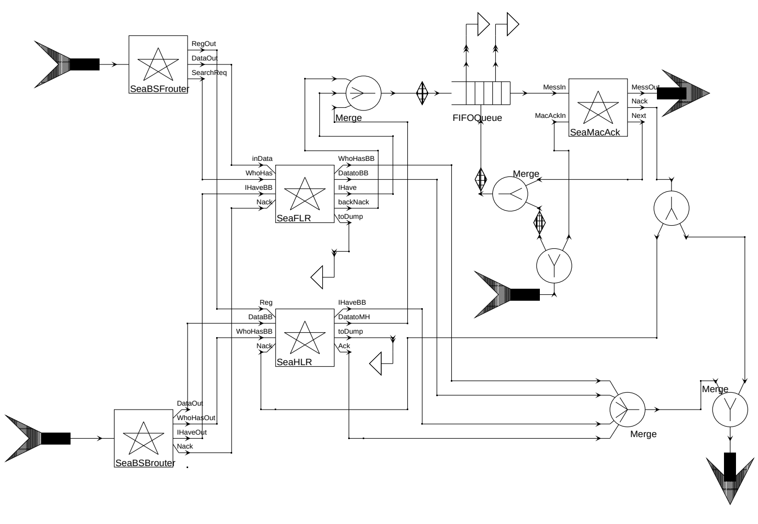


Figure 58:Model for Distribution Layer in Base Station, Reactive Search Scheme

The **multicast scheme** (Figure 59 and Figure 60) is integrated as an extension to the periodic search scheme. Compared to this scheme presented in Section 4.4., no changes need to be made to the functionality of the mobile host. However each base station is assigned a group of cells, that are geographically located around its location. Each packet destined to a host associated with this base station is also transmitted to the other members of the group of base stations and is temporarily stored there. In order to be able to incorporate this functionality into the simulation model, the HLR star is modified in a way that it receives not only all packets destined to mobile hosts it is currently serving but also all packets destined to hosts served by base stations that have the base station in its group. The packets destined to hosts that are associated in the local cell are sent - as before - over the wireless link to the receiver. The other, group-addressed packets, are stored in the newly added multicast buffer. The multicast buffer is constantly emptied from old data packets. Everytime a host becomes newly associated with the base station, it immediately signals to the HLR to checks its multicast buffer for packets that have already arrived for the newly associated host and transmits them immediately. These changes can be seen in Figure 59 with the additional signal path signalling a new registration to the star MC_Buffer and the data output path from this star to the star handling the output process.

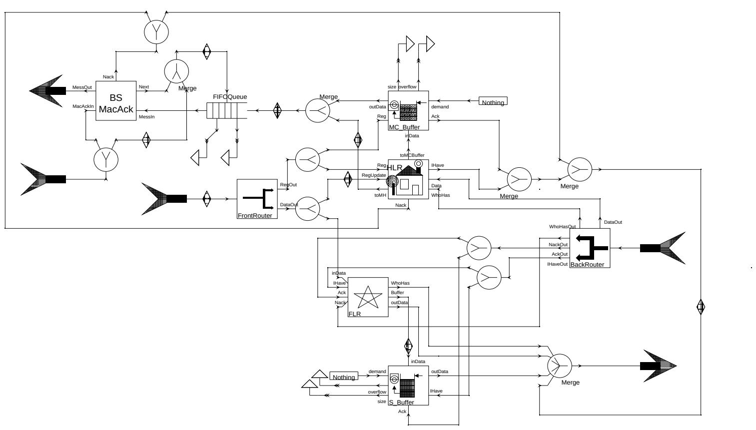


Figure 59:Model for Distribution Layer in Base Station, Multicast Scheme

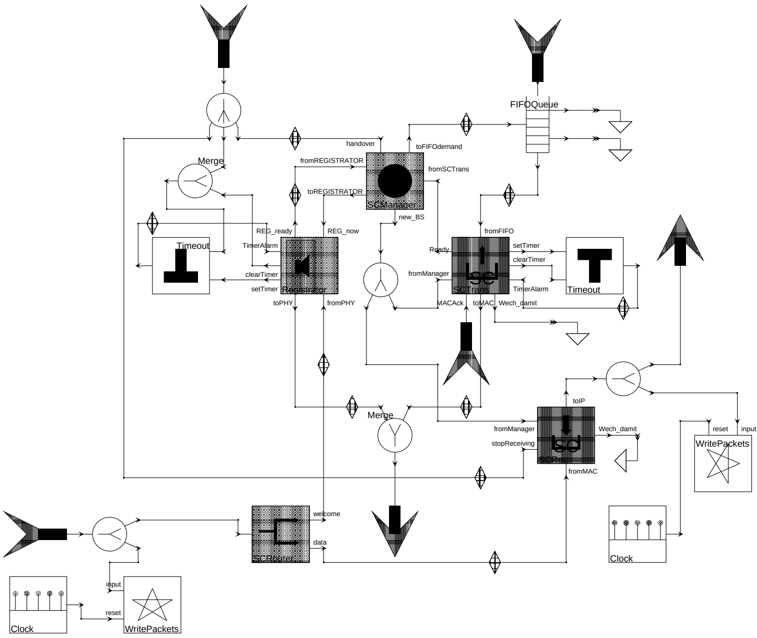


Figure 60:Model for Distribution Layer in Mobile Station, Multicast Scheme

Chapter 6

Performance Evaluation

6. Performance Evaluation

The following chapter will present the comparison and evaluation of the 4 schemes developed and described in the previous chapters - the periodic table scheme, the periodic search scheme, the reactive search scheme and the periodic search scheme improved with the multicast functionality. The evaluation is carried out with the help of the simulation model presented in the previous chapter. An extensive discussion of the simulation results is presented in this chapter.

6.1. Simulation Parameter:

In order to get comparable results the same setup for all schemes as shown in Figure 43 is simulated, i.e. 10 base stations serving 10 cells with 2 hosts. One of the mobile hosts remains stationary and sends towards the second mobile host that moves around through all the cells available. The second host itself is not transmitting itself unless acknowledgments or other control packets (registration) need to be sent upstream. This setup has been selected since it appears to be eliminating influence from factors other than those directly related to the performance of the schemes. One would not be able to gain clear results if multiple hosts generate unpredictable amount of background traffic thus causing random disturbances. Therefore the access delay has been set to a fixed, typical value (0.0001 s). Also it is not interesting to find out that the bandwidth of the distribution medium is too limited in order to transmit all the packets necessary. Therefore this bandwidth has been set to five times the available bandwidth of the wireless cells. The data users in the system were modeled to generate packets according to a Poisson process. The average generation rate λ was modeled over the set of values from 2 packets (each 500 byte long) per second (which corresponds to the rate of 8 kBit/s) to 20 packets/second (80 kbit/s), to 200 packets/second (800 kbit/s), to 1000 packets/second (4 Mbit/s), to 2000 packets/second (8 Mbit/s) and up to 4000 packets/second (16 Mbit/s).

The properties of wireless channels differ significantly from wired transmission mediums in terms of path loss, attenuation, fading, propagation parameters, bit error rates, collisions, etc. These effects result e.g. in a higher rate of packet losses / packet corruptions. For the evaluation of the distribution system however these effects are not relevant, the distribution system only intends to provide mobility management between isolated communication islands of any type, independent of the communication properties found in these islands. The simulation model used for the wireless channel therefore does not contain many of the properties of a wireless channel, in particular there are no packet losses or corruptions. This simplification is justified since the underlying wireless technology may be one of several with different characteristics, therefore *typical* properties can not be assumed. Of course this abstraction does not affect packet losses caused by handover events, that are one of the main points of interest in this evaluation.

The system does however provide a certain handover latency period (0.1 s) as it is likely to be found independent of what type of communication island is applied. Generally, if not stated otherwise, the handover frequency has been set to one handover in average every 10 seconds. This reflects a speed of 3.6 km/h for a cell diameter of 10 meters, 18 km/h for a cell diameter of 50 meters. For a handover frequency of one every 2 seconds the corresponding speeds translate to 10 km/h and 90 km/h. This seems to cover the speed as it can generally be expected in an in-door office environment. Further parameters can be found in the table below, however not all parameter apply to all models, some parameters are changed in some simulation runs which is stated there explicitly.

Parameter	Value
handover latency	0.1 s
search timer	0,1 s
MAC ACK timer	0.4 s
registration timer	4.0 s
access delay	0.0001 s
distribution medium bitrate	10 MBit/s
wireless bitrate	2 MBit/s
data packet length	500 byte
control packet length	64 byte
outdated time	3.0 s
multicast buffer	100 packets
new_search time	3.0 s
table frequency	1.0 s
handover rate	10 s

Table 1: Simulation Parameter

6.2. Simulation Results

The following section will first contain a presentation of a general comparison of the four schemes under evaluation, followed by an in-depth look into the dependencies of the separate schemes.

In the first subsection 6.2.1. the four approaches are compared against each other under comparable configurations, with a look at the delay characteristics and the throughput characteristics, each with increasing handover rates. Furthermore the average duration of the interruption caused by a handover event is compared between the four schemes as well as the overhead caused by the application of the four schemes in all the sections of the transmission path between the sender and the receiver - the first wireless section between the sender and the first base station, the section between the two base stations over the distribution medium and the section between the second base station and the receiver.

The following sections 6.2.2. to 6.2.5. present simulation results for each of the four schemes (the periodic table scheme in Section 6.2.2., the periodic search scheme in Section 6.2.3., the reactive search scheme in Section 6.2.4., the multicast-extended periodic search scheme in Section 6.2.5.), looking into the sensibility of the performance of the schemes towards changes of key parameters. Parallel to the general comparison in Section 6.2.1., the performance evaluation is carried out with respect to delay, throughput, interruption duration and bandwidth efficiency factor. Additionally the ratio of lost packets is looked at. The last subsection in this chapter wraps up the performance results.

6.2.1.General Comparison.

Figure 61 presents the simulated average delay curves over load for all four schemes at an handover frequency of 10s. As one can see the experienced **average** delay is almost identical for all four schemes. Under high load one experiences an average delay in the area of 2.3 to 2.5 ms. This rather similar result for the different schemes stems from the fact that the handover events are not too frequent in comparison to the packet transmission events, so that the few packets that are affected by the handover do not account for much change in the average figures.

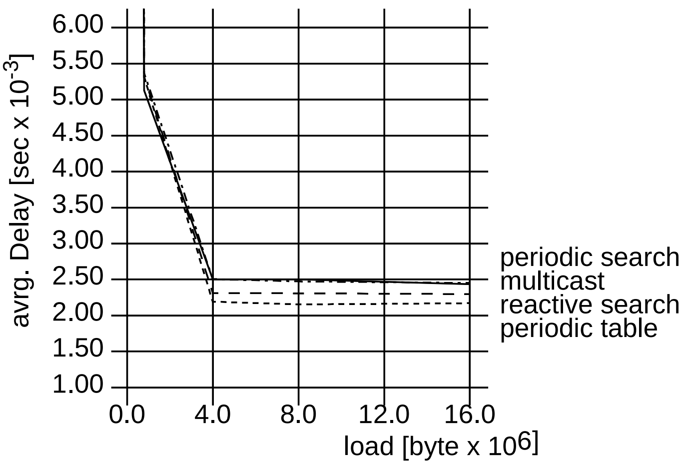


Figure 61:Average Delay over Load for all 4 Schemes, Handover rate 10s

The same observation can also be made if the handover frequency is double as high, i.e. one handover in average every 5 seconds (Figure 62). This figure gives the information that the ranking of the schemes against each other remains unchanged for increasing handover rates, just the absolute values decrease. This means that no scheme reacts worse than the others on the increasing handovers.

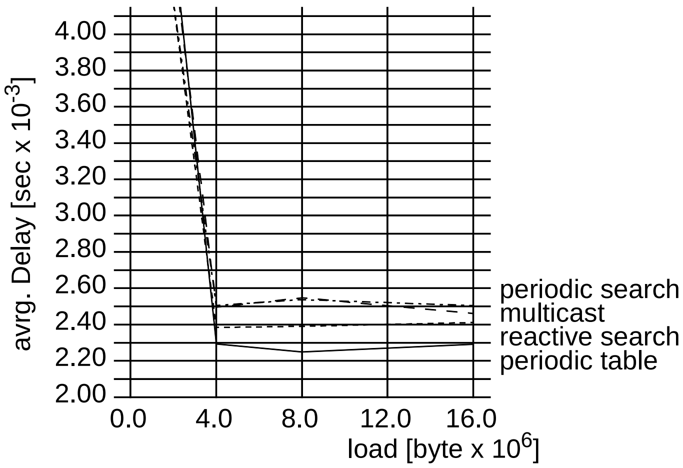


Figure 62:Average Delay over Load for all 4 Schemes, Handover Rate 5s

Parallel to the observations made in the delay figures, only small differences can be found when looking at the simulated accumulated throughput, that can be achieved with the four different schemes. Figure 63 shows the accumulated throughput for the 4 schemes at an average handover rate of 10s, Figure 64 shows the same figure for an handover rate of 5s.

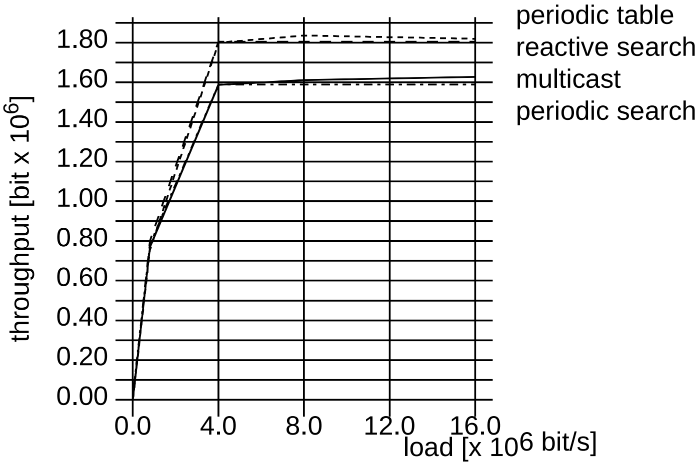


Figure 63:Accumulated Throughput over Load, all 4 Schemes, Handover Rate 10s

In both figures the curves do not differ significantly. The decrease in throughput in Figure 63 and Figure 64 for an increasing handover frequency is caused by the double as many accumulated interruption periods, when no packet can be delivered. Similarly to the delay figures the ranking of the schemes against each other remains identical, i.e. the changes of the performance of the schemes are identical. Not surprisingly a close relation between average delay and accumulated throughput can be seen in Figure 61 to Figure 64 - the lower average delay is seen in a simulation, the higher the accumulated throughput gets.

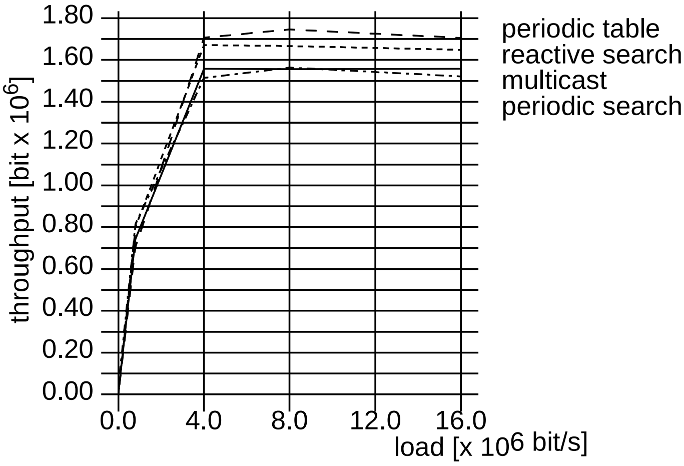


Figure 64:Accumulated Throughput over Load, all 4 Schemes, Handover Rate 5s

The previously discussed figures also show the positive effect of the added NACK-propagation feature. One might have expected that the two search schemes perform rather identically or rather with advantages for the periodic search scheme, since the search process has to be carried over one fewer hop. The graphs however show slightly better performance for the reactive search scheme, caused by the added NACK-propagation feature, only integrated in the reactive search scheme. Where the base station initiates a search only timer based when the information in its table has turned stale, the reactive search scheme makes use of the immediate notification of a failed transmission that is passed on across the distribution medium. This scheme could also be applied in the periodic search scheme which again would likely result in similar performance of both search schemes.

The relative success of the periodic table scheme compared to the other schemes in the 4 previous figures is related to the higher degree of simplicity in the scheme. Once the correct information of the receivers location is acquired, the scheme does not require any control packets to be transmitted. Therefore the phases **between** the handover events are used more efficiently and the possibly longer interruption times do not completely use up this advantage.

The differences between the four schemes become much more obvious when one only looks at the performance during the handover events itself - **the** moment of interest. The mean interruption period duration, i.e. the time that passes between the moment of the handover, when the mobile receiver looses connection to the old base station, and the moment of the arrival of the first data packet after the handover is shown in Figure 65. To give more precise figures, independent of the characteristics related to the physical handover (that is highly dependent on the implementation of the underlying physical layer), the duration of the handover latency has been subtracted from this interruption duration. See Figure 41 for information on the elements of the handover latency period.

As can be seen in the below figure the periodic table scheme, the periodic search scheme and the reactive search scheme perform almost identically in terms of mean interruption duration for their standard parameter setting (see Table 1, Simulation Parameter) in the area of 400ms. The simulation parameters have been set purposely to values that resulted in similar curves for the interruption duration in order to get a comparable reference point for the simulation parameter variations described later in the sections dealing with parameter variations on the performance of the system. The curve for the multicast scheme however can hardly be found in the figure since its mean interruption duration is about two magnitudes smaller, at around 3ms. Later it will be shown that this scheme not only is able to provide by far the shortest interruption times after a handover has occurred, but also provides them constantly and unaffected by changes in load condition or handover frequency, thus giving predictable behavior during a handover phase. This predictability is making this scheme especially suitable for multimedia communications with their unique requirements regarding quality of service guarantees.

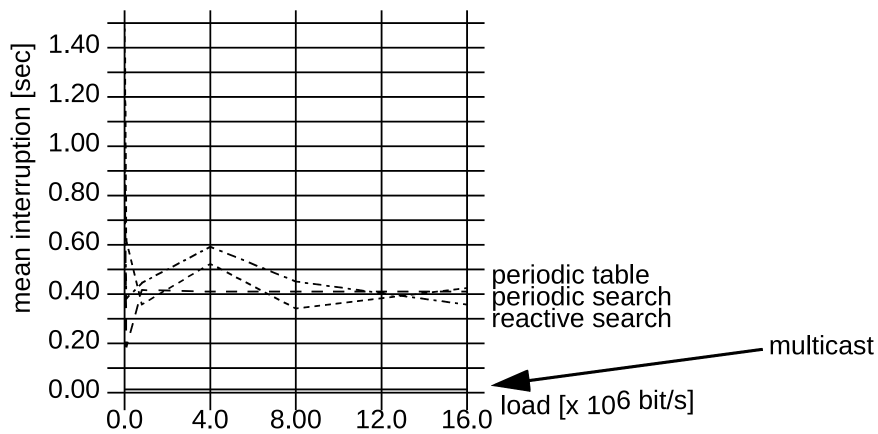


Figure 65:Mean Interruption Period Duration, all 4 Schemes, Handover Rate 10s

Another figure that makes the differences in the four schemes very clear can be got when one looks at the bandwidth efficiency factor in the distribution medium. Figure 66 presents the bandwidth efficiency factor - the ratio of data packets to control packets - in the distribution medium. A higher value in this figure translates to a higher efficiency, since more data packets can be transmitted per each control packet. The figure does not reflect the different size of data packets and control packets but merely counts each packet type. In the below figure one can see the advantage gained from the negative acknowledgment feature in the search scheme as described above. With the few search processes that get initialized only when actually a handover occurred the reactive search scheme provides by far the highest bandwidth efficiency factor on the distribution medium. Compared to this the periodic search scheme, due to its periodically applied searches, transmits 4 times more control packets in the same time. Since the multicast scheme is built on top of this scheme it performs almost identically in this figure. Quite obviously the periodic table scheme causes a high degree of overhead on the distribution medium with its continuous transmission of connectivity tables over the distribution medium resulting in the highest amount of control packets on the distribution medium.

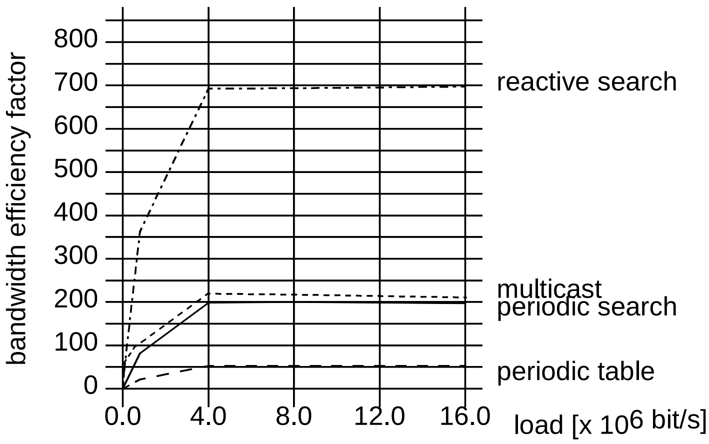


Figure 66:Bandwidth Efficiency Factor in Distribution Medium, all 4 Schemes, Handover Rate 10s

It has also to be noted that this overhead is independent of its actual necessity. This can be seen in Figure 67, where the bandwidth efficiency factor is shown again for all four schemes when no handover event occurred at all, i.e. in the case of a static sender **and** a static receiver (For better visibility a logarithmic scale on the y-axis has been used). Again the periodic table scheme produces the highest degree of overhead on the distribution medium, no matter of its actual necessity. Similarly the periodic searches initiated in the periodic search scheme and the multicast scheme are reducing the efficiency on the distribution medium and cause control overhead. The reactive search scheme performs extremely well if no handovers occur, since it never has to initiate any search at all. The curves in this figure again are highly dependent on the selected parameter setting like the table frequency and the search frequency, that can easily be altered with strong effect on the bandwidth efficiency factor, as will be shown later on. The parameter settings are set with respect to the reference point from Figure 65, that defines a setting, where the first three schemes perform with an almost identical mean interruption duration.

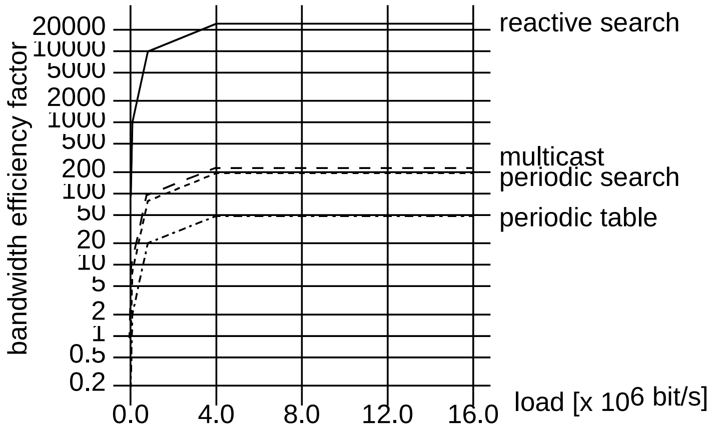


Figure 67:Logarithmic Bandwidth Efficiency Factor in Distribution Medium, all 4 Schemes, no Handover

The following Figure 68 shows the bandwidth efficiency factor on the sender side link.

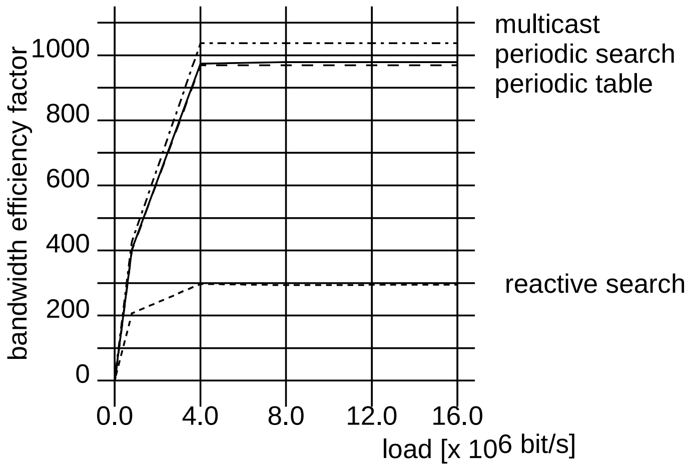


Figure 68:Bandwidth Efficiency Factor, Wireless Link on Sender Side, all 4 Schemes, Handover Rate 10s

The graph does not show differences in the overhead for neither the multicast scheme, nor the periodic search scheme and the periodic table scheme. Since those schemes do not incorporate any functionality between the sender and the initial base station there obviously is no reason for different overhead - only the reactive search scheme transmits additional packets (WHO_HAS, I_HAVE and NACK) over this link. The effect of the additional overhead caused by the sender-initiated searches in the reactive search scheme results in a much worse bandwidth efficiency factor in the figure below.Since one of the design criteria derived from the application scenario (see Figure 3 in Section 1.6.) was to have schemes with a passive receiver, all the four schemes do not involve exchanges other than data exchange and acknowledgments over the wireless link on the receiver side. Therefore the overhead on the wireless link on the receiver side of the path remains completely unaltered for all 4 schemes, which is why all 4 schemes have an identical bandwidth efficiency factor on this link. Due to this small information content the figure with these graphs have not been included here, however the figures comparing the bandwidth efficiency factor in each of the schemes (Figure 69, Figure 78, Figure 85, Figure 96) contain the curves for the receiver-side bandwidth efficiency factor.

6.2.2.Periodic Table Scheme

The following sections will present simulation results allowing a look into the characteristics of each of the four schemes as well as the dependencies they show towards alterations of critical parameters.

This subsection takes a look onto the periodic table scheme. The bandwidth efficiency factor will be compared as it is found in the 3 segments of the path between the sender and the receiver, i.e. the wireless link on the sender side, the link over the distribution medium and the wireless path on the receivers side. As shown in Figure 69 on both wireless links one observes almost the same bandwidth efficiency factor, only worsened on the receiver side by the control packets caused by the re-registration procedure.

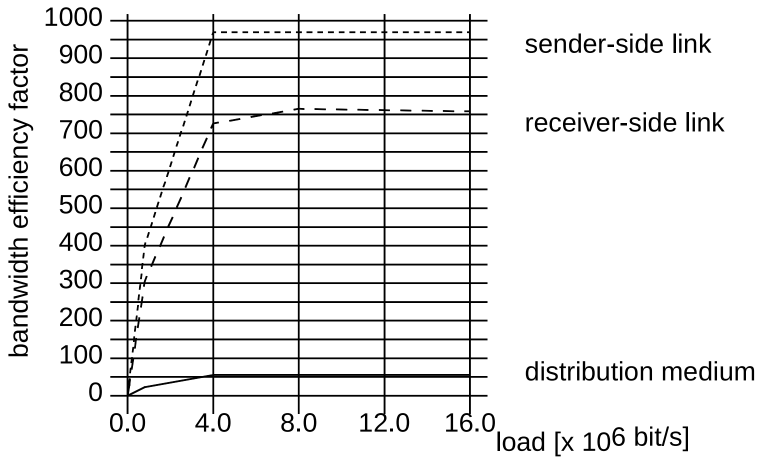


Figure 69:Bandwidth Efficiency Factor, Periodic Table Scheme, Handover Rate 10s

As it had to be expected the periodic transmissions of the connectivity tables causes a significant amount of overhead on the distribution medium. Almost 10 times as many control packets are transmitted over this link. This figure only represents the results for a table exchange rate of 1 s.The scheme is mostly unaffected by an increased handover frequency. Figure 70 shows that the mean interruption period duration for an increasing handover frequency (one per average every 10 seconds, 5 seconds and 2 seconds) remains mostly unchanged. Of course this observation is only valid as long as the table frequency is higher than the handover frequency. As the handover frequency comes closer to the table frequency the mobile system may remain unconnected for the full duration of its stay in a cell and may have moved already further on before the updated (and then already useless) information may reach the sender. Here again the table frequency is set to one every second.

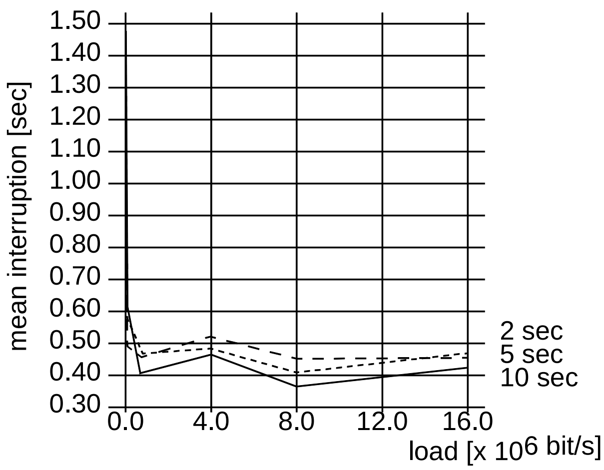


Figure 70:Mean Interruption Period Duration, Periodic Table Scheme, Handover Rate 10s, 5s, 2s

The figures for average delay (Figure 71) and accumulated throughput (Figure 72) for increasing handover frequency show some noticeable effect, even though the changes are not dramatically. The average delay increases from 2.16 ms (10s handover) to 2.27 ms (5s handover) and to 2.55 ms (2s handover).

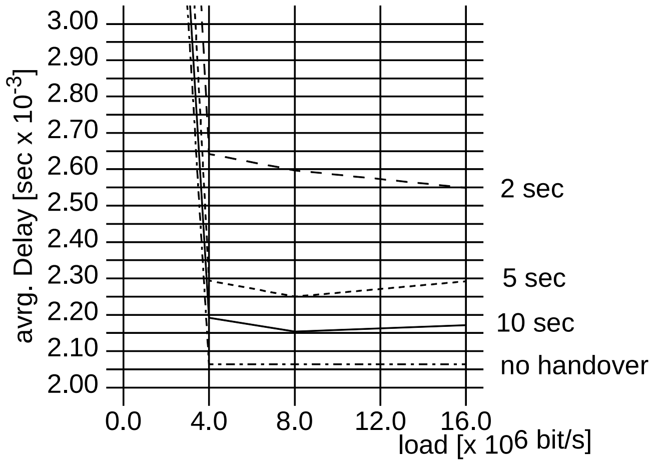


Figure 71:Average Delay, Periodic Table Scheme, Handover Rate none, 10s, 5s, 2s

Similarly the accumulated throughput goes down from 1.81 Mbit/s (10 s handover) to 1.70 Mbit/s (5 s handover) to 1.45 Mbit/s (2 s handover). For reference reasons both figures additionally contain a curve for the delay and throughput if **no** handover occurs.

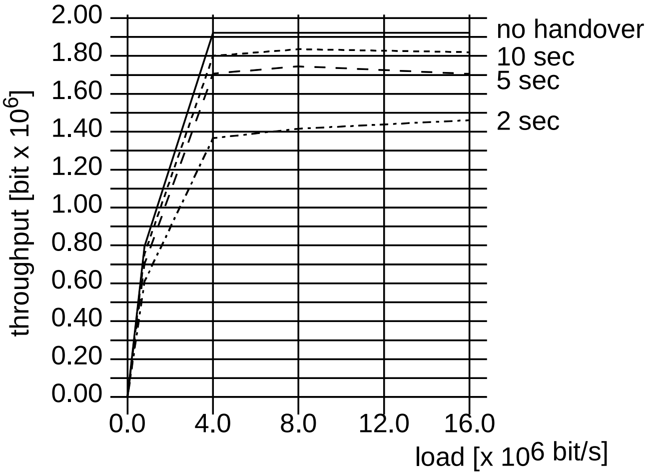


Figure 72:Accumulated Throughput, Periodic Table Scheme, Handover Rate none, 10s, 5s, 2s

In the following figures the system has been set to a handover frequency of one every 10 s again. The critical parameter in the periodic table scheme, that determines the performance is the frequency in which the tables are distributed among the base stations. In the following the results for different table exchange rates ranging from one every 4 seconds, 1 second to 0.4 seconds will be compared. Compared to the case with a table frequency set to 1s, one can see in Figure 73 and Figure 74 that the average delay actually does increase and the accumulated throughput does decrease if the tables are only sent every four seconds. The differences again are not very big due to the levelling impact of the average figures. Parallel one can see the opposite effect if the table frequency is set to one every 0.4 seconds.

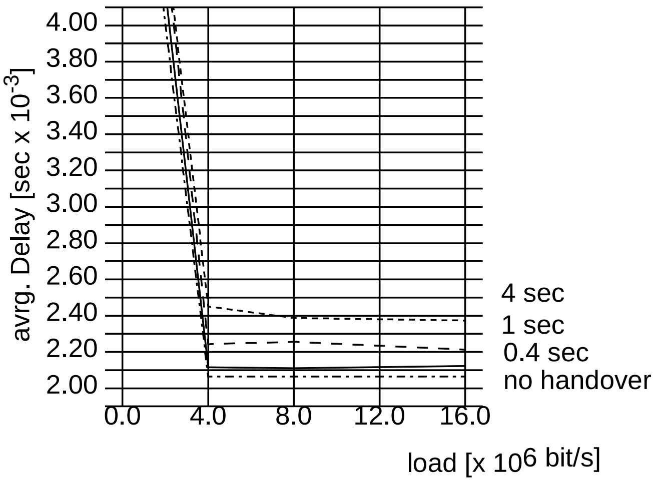


Figure 73:Average Delay, Periodic Table Scheme, Handover Rate 10s, Table Rate 4 s, 1 s, 0.4 s

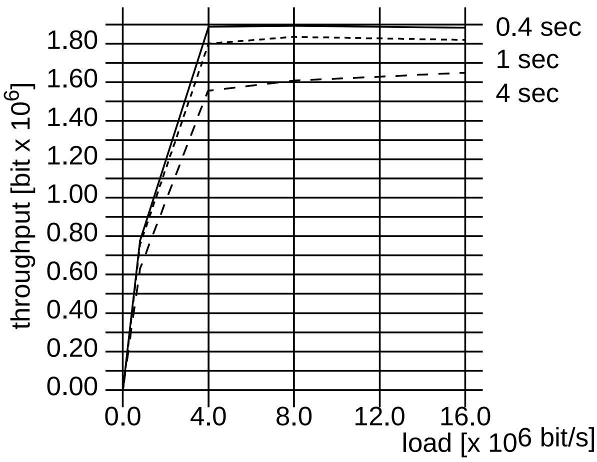


Figure 74:Accumulated Throughput, Periodic Table Scheme, Handover Rate 10s, Table Rate 4s, 1s, 0.4s

This effect - for worse, if the tables are sent less frequently, for better if the tables are sent more often - becomes a lot more apparent in the figures showing the mean interruption duration (Figure 75). For the slow table update strategy the average interruption duration exceeds far beyond one second, up to 1.6 seconds, whereas the fast update strategy approximately halves the interruption duration from 0.4 s to 0.2 s compared to the case for an update every second.

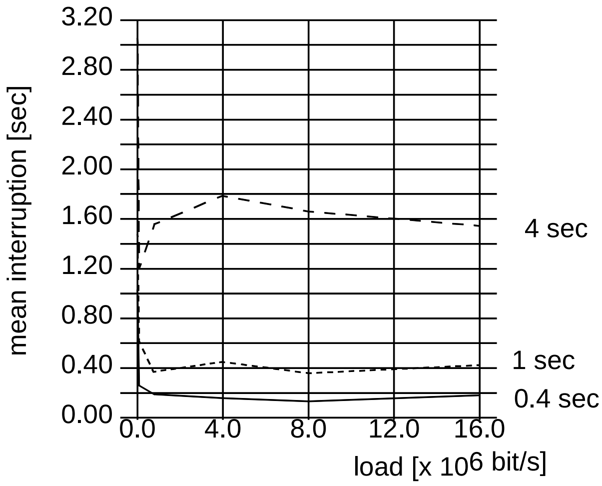


Figure 75:Mean Interruption Period Duration, Periodic Table Scheme, Handover Rate 10s, Table Rate 4s, 1s, 0.4s

Of course this improvement cannot be achieved at no cost. The more frequent tables have to be exchanged on the distribution medium the worse the bandwidth efficiency factor on the distribution medium gets (Figure 76). In the case with the fastest table update frequency one control packet has to be sent for every 20 data packets. If the tables are distributed ten times less often one can observe approximately ten times as many data packets per control packet on the distribution medium. However these many more data packets do not necessarily become successfully delivered to the receiver. If the mobile receiver has moved away during these 4 seconds the packets will be dropped at the end of the distribution medium.

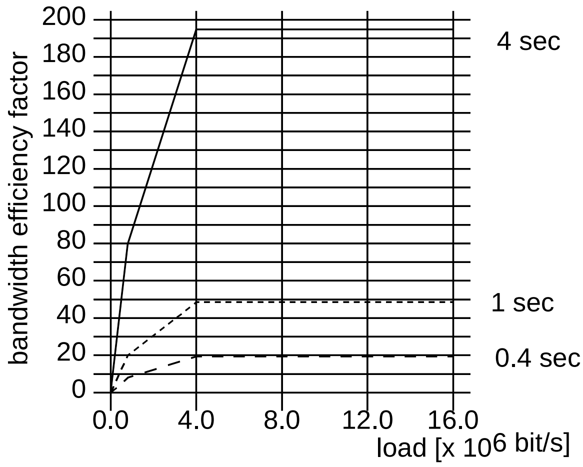


Figure 76:Bandwidth Efficiency Factor in Distribution Medium, Periodic Table Scheme, Handover Rate 10s, Table Rate 4s, 1s, 0.4s

This can be seen in Figure 77 showing the rate of lost packet, i.e. all data packets that get sent out but do not arrive at the destination, since they got dropped somewhere along the way. This rate of lost packets increases sharply for the slow table update frequency - almost every 5th packet does not arrive at the destination, compared to the other two update values, where only 6% or 3% respectively of the sent packets are lost.

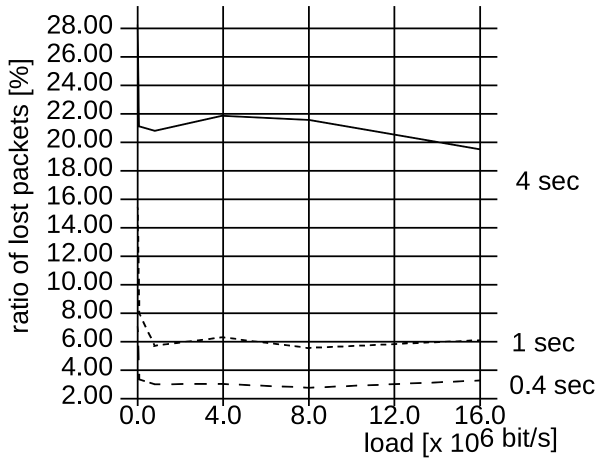


Figure 77:Ratio of Lost Packets, Periodic Table Scheme, Handover Rate 10s, Table Rate 4s,1s,0.4s

6.2.3.Periodic Search Scheme

Looking at the periodic search scheme one can see in Figure 78 that again the main control overhead occurs in the distribution medium compared to the two wireless links on the path. As before in the periodic table scheme the wireless link on the sender side has a higher degree of efficiency since no re-registration procedures have to be carried out.

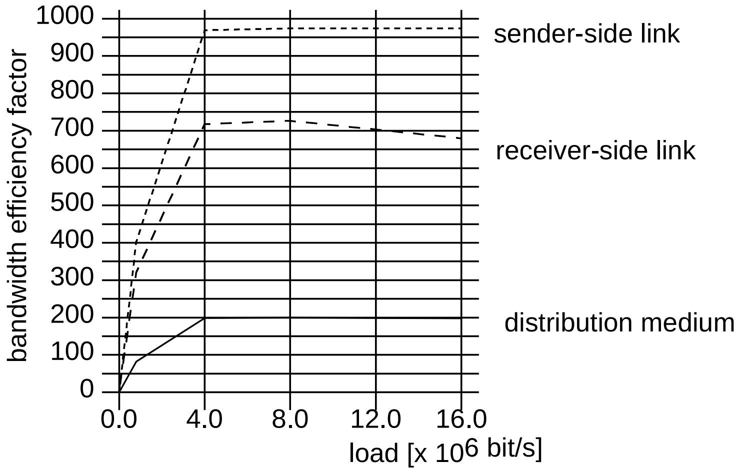


Figure 78:Bandwidth Efficiency Factor, Periodic Search Scheme, Handover Rate 10s

If the handover frequency is increased from 10 seconds to 5 seconds and to 2 seconds, one can see that the performance of the scheme, as shown in Figure 79 for the mean interruption period, Figure 80 for the average delay and in Figure 81 for the accumulated throughput, as before remains mostly unaffected. The interruption duration vary between 0.33 ms and 0.36 ms, the average delay increases at ~0.25 ms from 2.5 ms to 2.75 ms. Slightly better visible are the differences in the throughput figure mainly again due to the fact that more interruption periods accumulate.

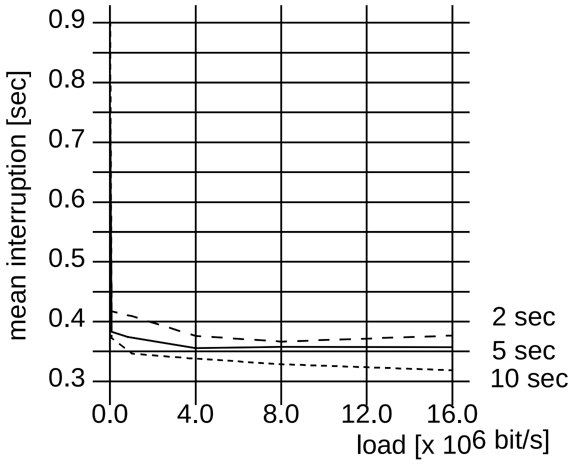


Figure 79:Mean Interruption Period Duration, Periodic Search Scheme, Handover Rate 10s, 5s, 2s

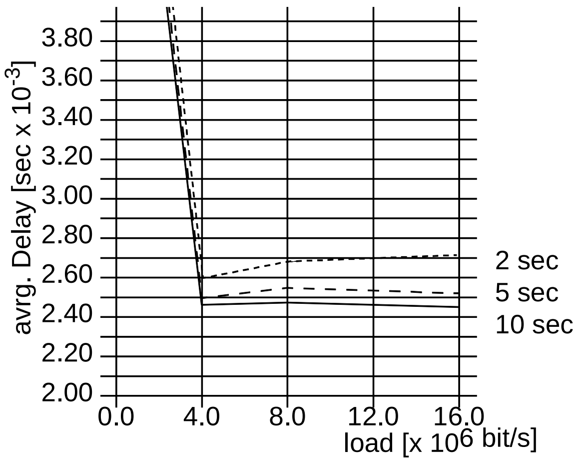


Figure 80:Average Delay, Periodic Search Scheme, Handover Rate 10s, 5s, 2s

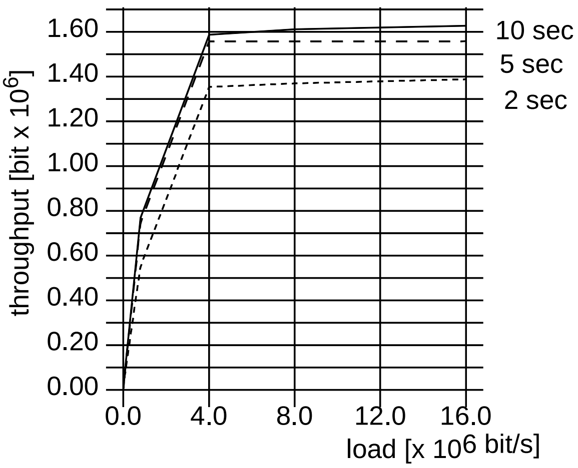


Figure 81:Accumulated Throughput, Periodic Search Scheme, Handover Rate 10s, 5s, 2s

The critical parameter with most significant influence in determining the performance of this scheme is the frequency in which periodically old entries are staled and new searches are initiated. For all previous figures on the periodic search scheme (Figure 78 to Figure 81) a search frequency of one second has been used. For the following figures the curves for a search frequency of 2 seconds and 0.1 seconds, each at a handover frequency of 10 seconds have been added. The mean interruption duration reflects the effect of the changes of this parameter most clearly (Figure 82). The mean interruption period lasts almost a full second, if the search frequency is increased to 2 seconds, it decreases dramatically if a search is carried out after every 0.1 seconds.

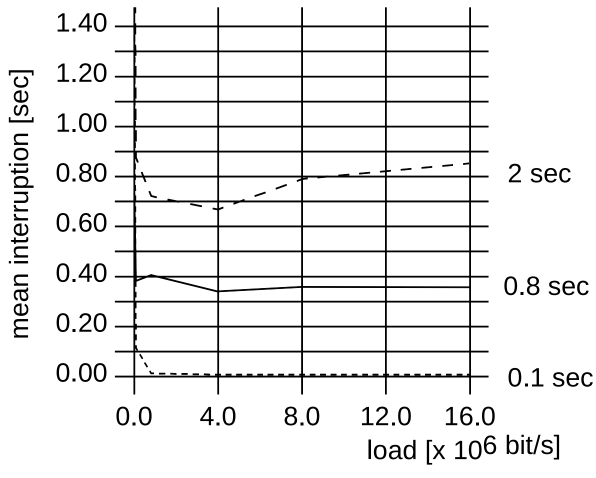


Figure 82:Mean Interruption Period Duration, Periodic Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.1s

The more frequent searches may not be needed in most times however, when the mobile host did not cross a cell boundary. Therefore the huge increase in the overhead on the distribution medium (Figure 83) caused by the frequent searches may not be justified, especially for a mobility as it is assumed in this environment.

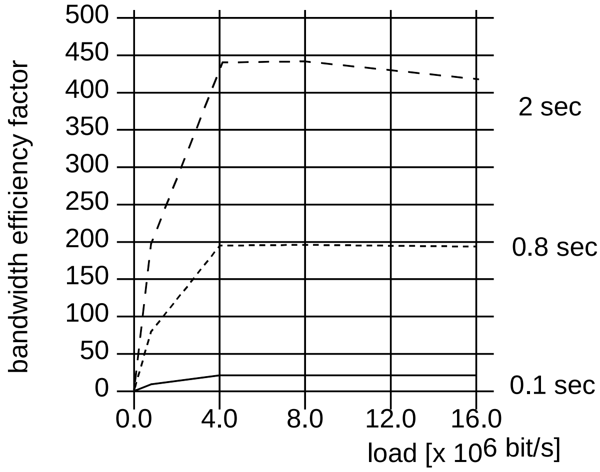


Figure 83:Bandwidth Efficiency Factor in Distribution Medium, Periodic Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.1s

The rate of lost packets increases as the search frequency decreases (Figure 84) from below 2% (search rate 0.1 s) to almost 4% (search rate 0.8 s) up to 6% (search rate 2 s), as it could be expected.

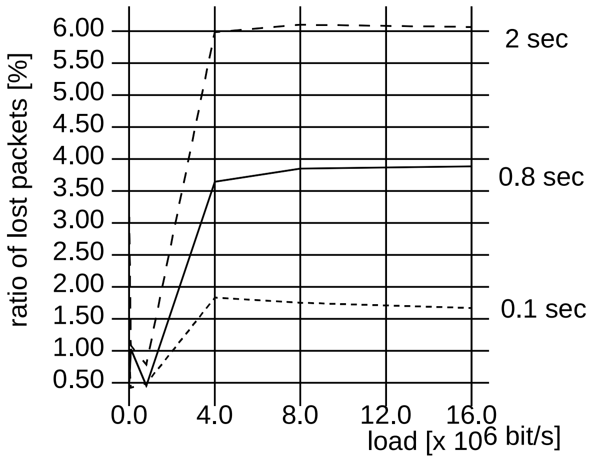


Figure 84:Ratio of Lost Packets, Periodic Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.1s

6.2.4.Reactive Search Scheme

Looking at the bandwidth efficiency factor in the reactive search scheme (Figure 85) it is no surprise to find that the highest overhead is no longer found in the distribution medium as in the previous schemes but in the wireless link at the sender. The searches, that have to be initiated at the sender cause a drop of the bandwidth efficiency.

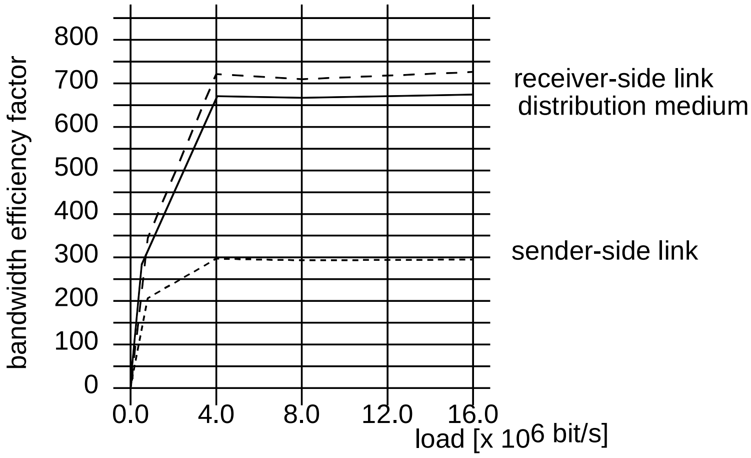


Figure 85:Bandwidth Efficiency Factor, Reactive Search Scheme, Handover Rate 10s

If the handover rate is increased from 10 seconds to 5 seconds and to 2 seconds one again observes only small influence on the mean interruption duration (Figure 86), average delay (Figure 87) and accumulated throughput (Figure 88).

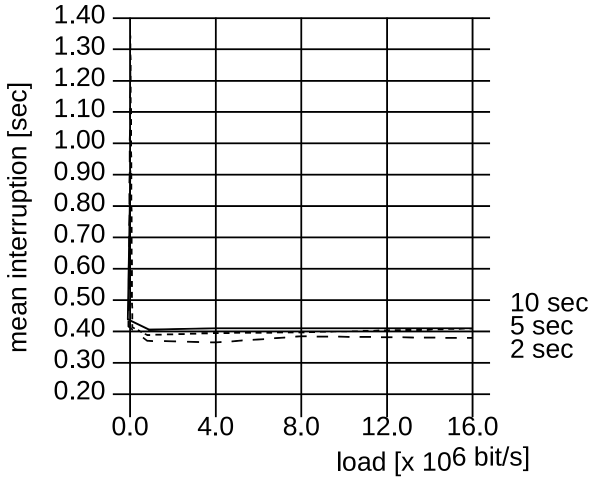


Figure 86:Mean Interruption Period Duration, Reactive Search Scheme, Handover Rate 10s, 5s, 2s

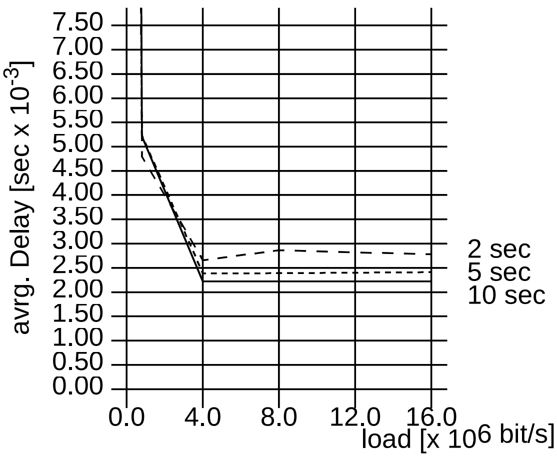


Figure 87:Average Delay, Reactive Search Scheme, Handover Rate 10s, 5s, 2s

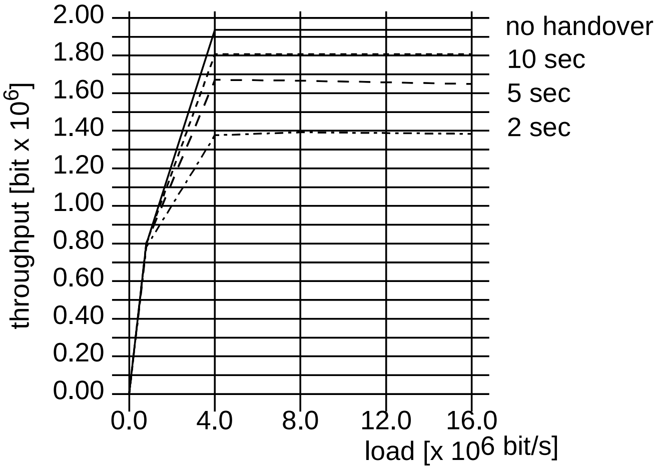


Figure 88:Accumulated Throughput, Reactive Search Scheme, Handover Rate none, 10s, 5s, 2s

As in the previously discussed periodic search scheme the crucial parameter with high effect on the systems performance in the case of handover is the time that determines how fast a new search is initiated. This parameter has been set to 0.4 s in the simulations so far. For the following figures the curves resulting from setting the search timer to 2 seconds, 0.8 seconds and 0.1 seconds, all at a handover rate of 10 seconds are also presented. Figure 89 shows that the average delay can be reduced up to a certain degree, as well as the accumulated throughput that increases with a smaller search parameter (Figure 90).

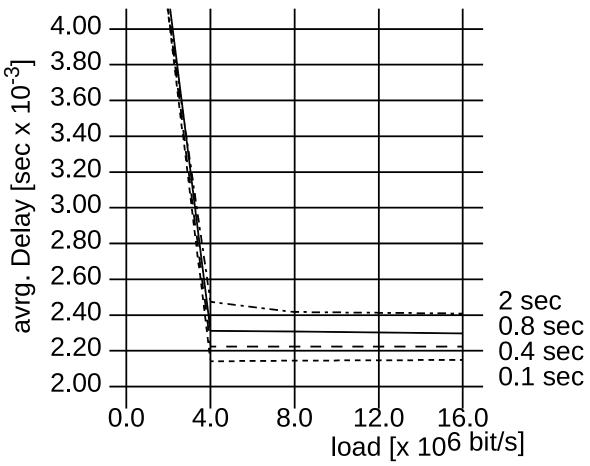


Figure 89:Average Delay, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.4s, 0.1s

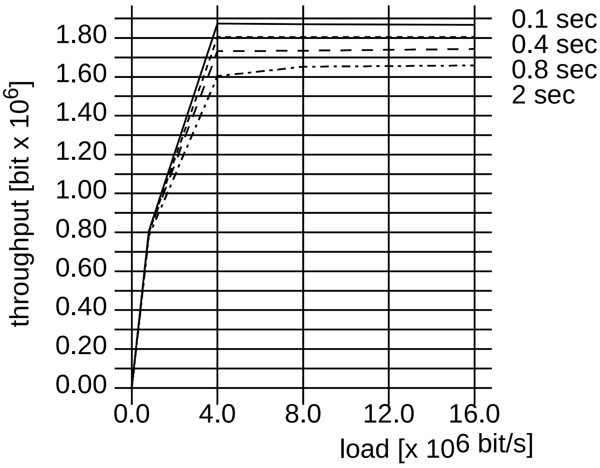


Figure 90:Accumulated Throughput, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.4s, 0.1s

The best indicator showing the effect of a faster search initiation can again be seen in the mean interruption duration (Figure 91). If the search timer expires only after 2 seconds the average length of the interruptions rises up to ~1 second. For very short search timer the average interruption duration can be reduced dramatically to the area of 0.01 seconds.

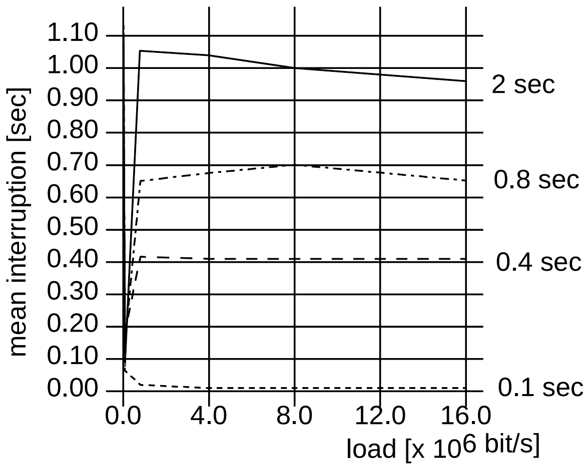


Figure 91:Mean Interruption Duration, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.4s, 0.1s

Surprisingly this does not have negative impact neither on the distribution medium (Figure 92) nor on the wireless links (Figure 92 for the sender-side wireless link, Figure 94 for the receiver-side wireless link). The reason for this stable behavior is once again the fact that - unlike in the model for the periodic search scheme - the previously explained NACK propagation scheme is applied, that reduces the number of unnecessary searches.

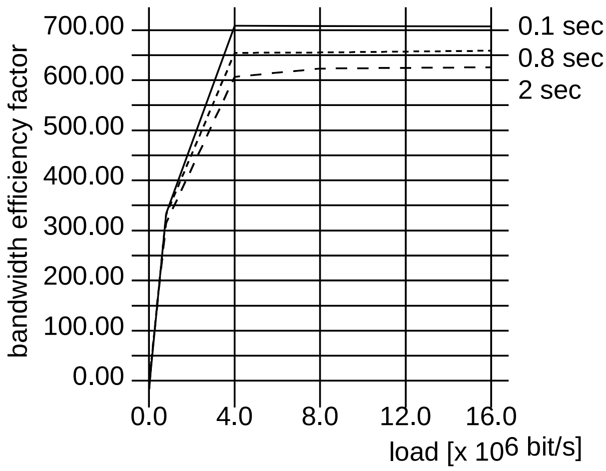


Figure 92:Bandwidth Efficiency Factor, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.1s

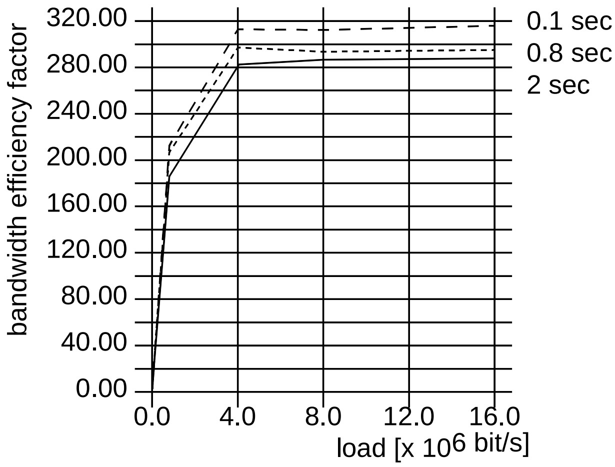


Figure 93:Bandwidth Efficiency Factor on the Sender-Side Wireless Link, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.1s

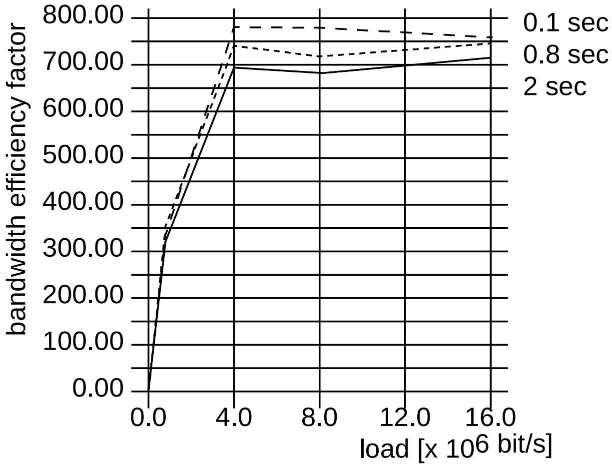


Figure 94:Bandwidth Efficiency Factor on the Receiver-Side Wireless Link, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.1s

Parallel to the above observed unaffectedness on the bandwidth efficiency factor one can also see almost no influence on lost packets ratio, as shown in Figure 95. This may be surprising in the beginning, however it becomes clear when one notices that the sender automatically stops sending as soon as a negative acknowledgment has arrived. This source blocking cannot be achieved in the other schemes, where the source is not informed about the interruption and therefore continues sending during the handover.

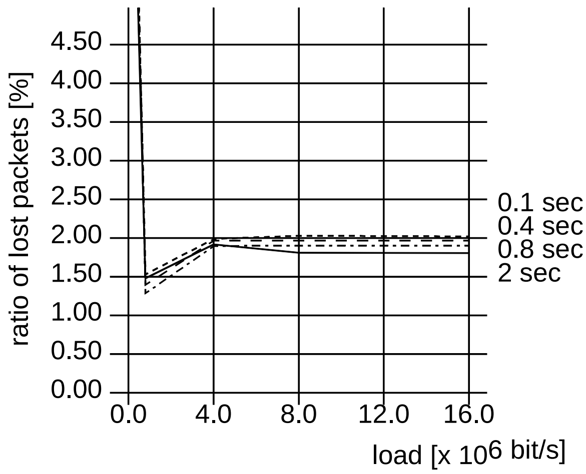


Figure 95:Ratio of Lost Packets, Reactive Search Scheme, Handover Rate 10s, Search Timer 2s, 0.8s, 0.4s, 0.1s

6.2.5.Multicast Scheme

As explained in Section 4.7. the multicast scheme is an improvement of the periodic search scheme. While the periodic search scheme (as the other two schemes) is used to provide the base stations with the location information used to create an up-to-date routing table, the multicast extension helps to shorten the interruption time. Therefore the bandwidth efficiency factor of the multicast scheme is almost the same as in the periodic search scheme (Figure 96).

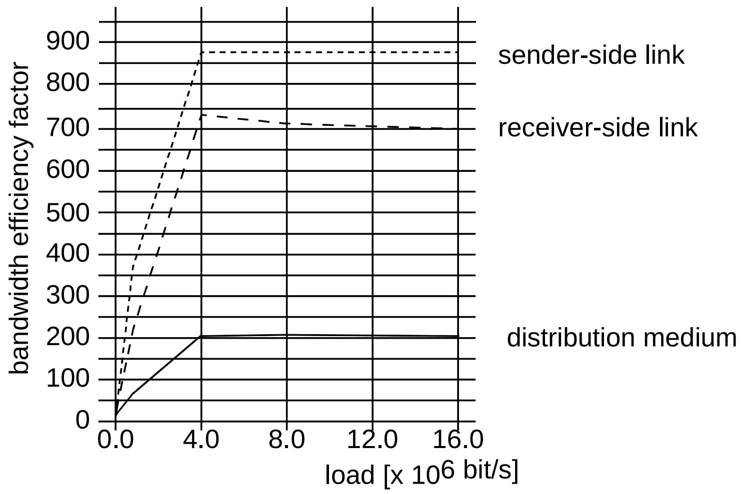


Figure 96:Bandwidth Efficiency Factor, Multicast Scheme

Note that this is only true in networks where true multicast is available, i.e. a packet that is sent to a multicast address is in fact transmitted only once and received by all members of the multicast group. If the multicast transmission is implemented in a way that separate point-to-point transmissions are taking place in order to reach each member of the multicast group, the overhead in the distribution medium multiplies respectively. The main overhead is located in the distribution medium, the wireless links have significantly less overhead to carry, where the receiver-side wireless link additionally has to transmit registration and registration reply packets.

The main motivation to invest the buffer space for the multicast packets in all base stations, to manage the multicast groups and deal with multicast transmissions is the expected significant reduction in the mean interruption period duration. This expectation is proved to be justified if one looks at the results from the simulations shown in Figure 97. The mean interruption duration remains unchanged, independent of the load or of the frequency in which a handover occurs. In absolute numbers the average duration is by far shorter than the interruptions experienced in the other schemes.

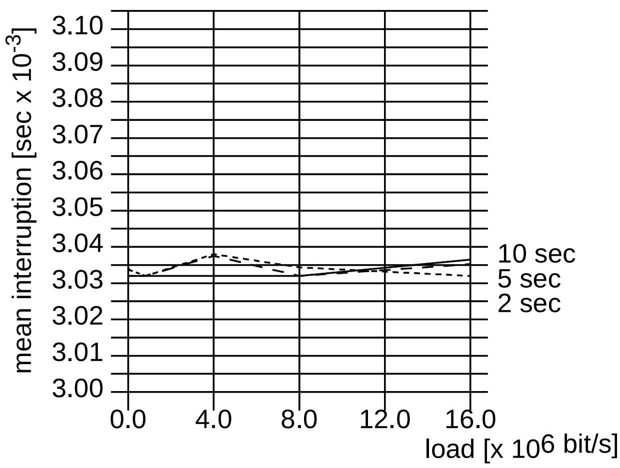


Figure 97:Mean Interruption Period Duration, Multicast Scheme, Handover Rate 10s, 5s, 2s

This was to be expected since the scheme is able to continue transmitting data packets immediately after the reregistration between the mobile host and its new base station is finished (see Figure 41 for an illustration of the achievable gain with the multicast scheme). Since this reregistration packet is likely to only consist of two packets - one request upstream, one confirmation downstream - that are only transmitted over one hop, chances for a disturbing influence, that may prolong the interruption phase are very low

With the mean interruption duration being unaffected by an increased handover frequency, it is no surprise to find that also the average delay (Figure 98) and the accumulated throughput (Figure 99) are almost not influenced by this as well.

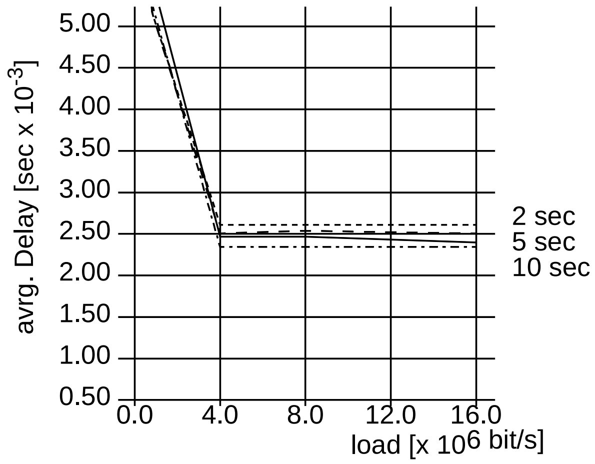


Figure 98:Average Delay, Multicast Scheme, Handover Rate 10s, 5s, 2s

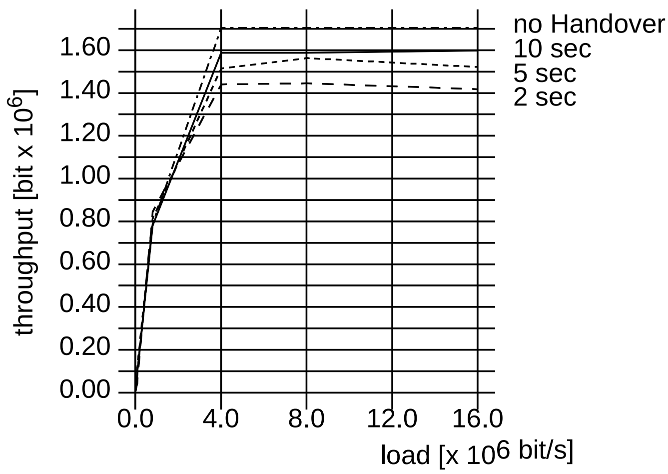


Figure 99:Accumulated Throughput, Multicast Scheme, Handover Rate none, 10s, 5s, 2s

Only the ratio of lost packets increases slightly (Figure 100) due to the fact that packets may have been removed from the buffer already, whose size has been set small to avoid duplications caused by transmissions in the old cell **and** in the new cell or by the duplications, that are counted to the lost packet ratio.

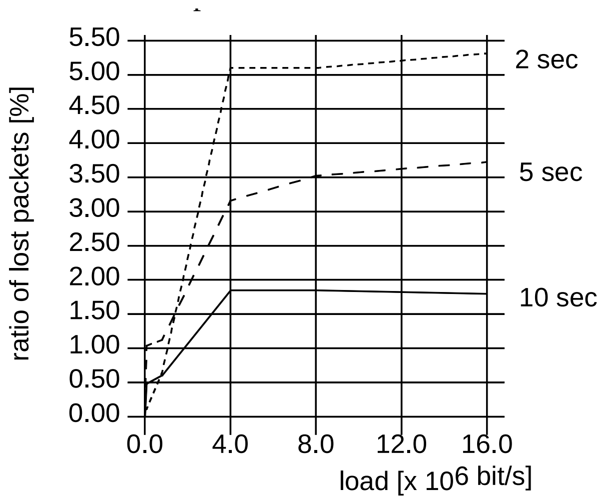


Figure 100:Ratio of Lost Packets, Multicast Scheme, Handover Rate 10s, 5s, 2s

If the curves of the interruption period duration for the very fast search schemes as well as for the periodic table scheme with very frequent table distribution are compared with the interruption time of the multicast scheme (Figure 101, only the search based schemes and multicast scheme in Figure 102) one can easily see that still the achievable interruption duration are significant longer for the first three schemes compared to the multicast scheme. The periodic table scheme does not go below 0.1s, but also the best results for the two search based schemes (0.01s and 0.008s) are not as good as for the multicast scheme, that performs stable at 0.003s.,

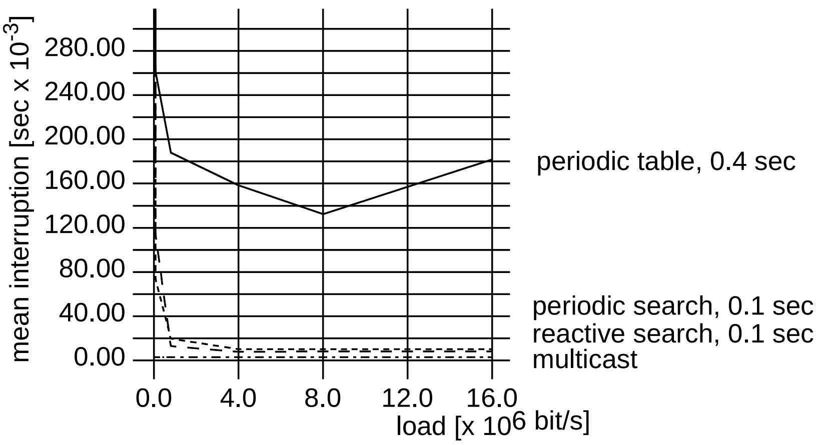


Figure 101:Mean Interruption Duration, Periodic Table Scheme w. Table Rate 0.4s, Periodic Search Scheme w. Search Rate 0.1s, Reactive Search Scheme w. Search Timer 0.1s, Multicast Scheme

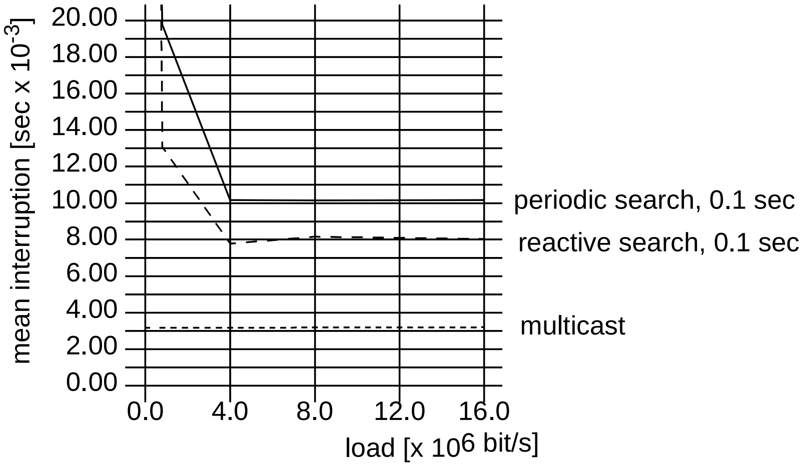


Figure 102:Mean Interruption Duration, Periodic Search Scheme w. Search Rate 0.1s, Reactive Search Scheme w. Search Timer 0.1s, Multicast Scheme

6.2.6.Performance Evaluation Wrap-up

With the help of the presented performance evaluation it has been shown that the **average** delay and the **accumulated** throughput of all four schemes do not differ much neither between the schemes itself nor with varying handover frequency nor with variations of sensible parameters in each scheme. The schemes equally degrade/improve with increased/decreased handover frequency, none of the schemes shows significantly different reaction on the change of the handover rate. Significantly different performance however can be seen if the performance in the most critical moments - the handover moments itself - is compared. The periodic table scheme and the two search based schemes all achieve similar **interruption durations** in the area of 0.4 seconds[1] for general parameter settings. With increased update or search effort this interruption duration can be reduced in those three schemes up to a certain degree, and of course be worsened with decreasing effort. The periodic table scheme degrades worse and improves less far compared to the search based schemes whose range of operation is higher.

The two search based schemes would usually perform almost identically, since the single hop, that the search is carried on additionally would account for much performance difference in terms of interruption duration. The difference would however be noticeable in the additional use of the wireless link. The choice for either of the two search schemes would therefore depend on syntactical issues, i.e.reactive search scheme does not allow for self learning through packet sniffing

1. Reminder: the duration period is defined the time between the arrival of the handover event at the mobile host until the first successful transmission of a data packet minus the handover latency time

on the back bone, the reactive search scheme used additional wireless bandwidth and uses additional power in the mobile sender, the reactive search scheme either requires cooperation of senders sending from outside the network into the E-WLAN or requires a special solution for this case, if cooperation cannot be expected. The periodic search scheme however requires the base station to listen to every packet on the wireless link - be it addressed to it or not - since the packets are sent out without any knowledge on the location of the destination, and decide on-the-fly, based upon its information in the registration table, whether the packet has to be forwarded over the distribution system or not. This requires some processing effort in the base station that adds to the processing delay there.

Both search schemes will benefit from the application of the NACK passing scheme since unnecessary searches can be avoided through this. This requires however that the MACACK timer in the base station, where the mobile host used to be connected to, that is responsible for the generation of the NACK to be sent upstream is faster than the search timer applied in the station receiving the NACK.

All the three schemes are capable of solving the mobility related situation at an acceptable speed, if the assumed communication characteristics are those of typical internet data traffic and if the mobility remains within certain bounds[1]. The additional effort that has to be invested in the multicast scheme (buffer space, management of multicast groups, multicast transmission) however is justified by the reduction of the mean interruption period duration of approximately one magnitude compared to the best results received from the search schemes. In addition to this shorter duration of the interruption, the performance of the scheme is widely unaffected by load increase or increased handover frequency. The system is capable to provide **predictable behavior** in the event of a handover. This predictability is a most valuable feature for the integration of the distribution system into a quality-of-service-providing infrastructure.

1. the targeted in-door, office type environment will likely not allow a much higher degree of mobility

Chapter 7

Conclusions

7. Conclusion

This section concludes this thesis. In the previous chapters solutions to build up a local area data communication infrastructure consisting of several small wireless communication cells have been developed. Based on an analysis of the application scenario an architecture, that is best suitable in that environment, has been chosen. The issues that need to be solved in setting up such an architecture have been identified. Followed by a discussion of alternate approaches, that may appear to be applicable as well, the architecture has been discussed with respect to its dependencies and effects on possible solution strategies. Issues concerning the design of the elements of the architecture - the mobile host, the base station, the distribution medium, the portal and shared resources connected directly to the distribution medium - are discussed and lead to the development of several distinct strategies that are applicable to provide the necessary functionality.

Four different approaches to solve the mobility management problem in the proposed architecture have been developped. The approaches have been specified and presented. The four approaches have been evaluated with the help of simulation, the results of the performance evaluation allow to compare the four approaches, to identify key parameter in the application of each of them, and to estimate the effort that is necessary to operate the schemes. Finally the issues in applying the architecture in a QoS-supporting infrastructure are discussed.

The results show that the duration of the interruptions occurring in the event of a handover between two cells can be influenced with the different schemes to a large degree. While the performance over a larger timeperiod does not show major differences, the performance in the critical event of the handover can be reduced or increased by parameter setting or by choosing a different scheme in the order of magnitudes. With appropriate parameter setting even the simplest scheme may provide sufficiently good service - depending on the requirements imposed by the configuration and the application. If very strict enforcement of tight QoS requirements is demanded one will have to select the multicast supported scheme, that is able to provide minimal interruption duration.

References

[1] IBM Software Glossary, http://www.networking.ibm.com/nsg/nsgabout.htm

[2] CISCO Systems list of Internetworking Terms and Acronyms, http://www.cisco.com/univercd/cc/td/doc/cisintwk/ita/index.htm

[3] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Draft Standard IEEE 802.11 , P802.11/D5.0, The editors of IEEE 802.11,; Jul 1996

[4] Hiperlan Functional Specification, Draft prETS 300 652, ETSI Secretariat,, Jul 1995

[5] Handoff in Hybrid Mobile Data Networks, K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. Makela, R. Pichna, J. Vallström, IEEE Personal Communications, April 2000

[6] Call for IEEE802.11 proposals on MAC enhancements and IAPP, John Fakatselis, IEEE 802.11 working group document IEEE 802.11-00/024; January 2000

[7] Wireless Data Networking, Nathan J. Muller, 1995, Artech House Boston, London

[8] Remaining Technical “NO” Comments on LMSC Ballot on IEEE P802.11 D5.0, Vic Hayes, IEEE 802.11 Chair, IEEE 802.11 working document 97/34

[9] How Far Should an ESS Extend?, Michael Fischer, IEEE 802.11 working group paper

[10] Wireless Base Station, Radio Equipment and Systems, Digital European Cordless Telecommunications, Draft prETS 300 175-10, European Telecommunications Standards Institute, Jan 1995, prStandard

[11] Mobilfunknetze und ihre Protokolle, Band 1 + 2, B. Walke, B.G.Teubner Verlag, Stuttgart, 1998

[12] Routing in Ad Hoc Networks, D. Johnson, Proceedings of the Workshop on Mobile Computing Systems and Applications, pp. 158-163, IEEE Computer Society, Santa Cruz, CA, Dec 1994

[13] Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations S. Corson, J. Macker, Internet Draft draft-ietf-manet-issues-00.txt, September 1997

[14] Architectural Considerations for Mobile Mesh Networking, S. Corson, S. Batsell, J. Macker, Internet Draft RFC Version 2, May 1996.

[15] A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks, Vincent D. Park and M. Scott Corson, Proceedings of IEEE INFOCOM '97, Kobe, Japan (April 1997)

[16] A Performance Comparison of TORA and Ideal Link State Routing, Vincent D. Park and M. Scott Corson, Submitted to IEEE INFOCOM '98, San Francisco, USA

[17] Ad Hoc On Demand Distance Vector (AODV) Routing, Charles Perkins, 20 November 1997, Internet Draft, draft-ietf-manet-aodv-00.txt

[18] The Zone Routing Protocol (ZRP) for Ad Hoc Networks, Zygmunt J. Haas, Marc R. Pearlman, November 1997 Internet Draft, draft-zone-routing-protocol-00.txt

[19] ANSI/IEEE Standard 802.1D - 1990, MAC Bridges, IEEE Project 802, Local and Metropolitan Area Networks, March 8, 1991

[20] Analyzing and Tuning the Distributed Coordination Function in the IEEE 802.11 DFWMAC Draft Standard, Jost Weinmiller, Hagen Woesner, Adam Wolisz , Proc. of the International Workshop on Modelling MASCOT '96, Analysis and Simulation of Computer and Telecommunication Systems; Feb. 96, San Jose, California

[21] Analyzing the RTS/CTS Mechanism in the DFWMAC Media Acces Protocol for Wireless LAN's, Jost Weinmiller, Hagen Woesner, Jean-Pierre Ebert, Adam Wolisz , Proc. of IFIP TC6 Workshop Personal Wireless Communications (Wireless Local Access) April '95 Prague, Czech Republic Jan Slavik, Otto Spaniol (editors) Aachener Beitraege zur Informatik, Band 9 Verlag der Augustinus Buchhandlung, 1995, pp. 117-130

[22] Performance Study of Access Control in Wireless LANs - IEEE 802.11 DFWMAC and ETSI RES10 HIPERLAN, J. Weinmiller, M. Schlaeger, A. Festag, A. Wolisz, in "Mobile Networks and Applications", Baltzer Science Publishers, July 1997 .

[23] ARP - Address Resolution Protocol, RFC 826, David C. Plummer, November 1992

[24] Wireless Access to High Speed Networks, A. Wolisz, J. Weinmiller, M. Schlaeger, H. Woesner, October 1995 in: High-Speed Networking for Multimedia Applications, W. Effelsberg, O. Spaniol, A. Danthine, D. Ferrari (eds.), Kluwer Academic Publishers, Boston/Dordrecht/London, 1996 .

[25] MACA - A new Channel Access Method for Packet Radio, P. Karn, ARRL/CRRL Amateur Radio 9th Computer Networking Conference, Sept. 22 1990

[26] MACAW: A Media Access Protocol for Wireless LAN‘s, V. Bharghavan, A. Demers, S. Shenker, L Zhang, SIGCOM 1994

[27] Solutions to Hidden Terminal Problems in Wireless Networks, C. L. Fullmer and J.J. Garcia-Luna-Aceves, Proc. ACM SIGCOMM 97, Cannes, France, September 14-18, 1997.

[28] A wireless LAN employing Distributed Radio Bridges; Victor C.M. Leung and Andrew W.Y.Au, in: Wireless Networks, Baltzer Science Publishers, vol. 2, no. 2, pp. 97-107, 1996

[29] IP Mobility Support, Internet Standard, RFC 2002, Ocktober 1996

[30] IP Encapsulation within IP, Internet Standard, RFC 2003, October 1996

[31] Minimal Encapsulation within IP, Internet Standard, RFC 2004, October 1996

[32] Route Optimization in Mobile IP Charles Perkins, David B. Johnson, Internet Draft, November 1997, draft-ietf-mobileip-optim-07.txt

[33] Implementation and Analyses of the Mobile-IP Protocol, F. Broman, F. Tarberg, Report on a degree project for a masters degree in Computer Science at KTH Stockholm, Jan. 96, http://www.it.kth.se/~d91-fta/exjobb/exjobb.html

[34] Fast and Scalable Handoffs for Wireless Internetworks; R. Caceres, V. Padmanabhan, Proc. of ACM MobiCom '96, November 96

[35] the Daedalus project, http://daedalus.cs.berkeley.edu,

[36] Handoffs in Cellular Wireless Networks: The Daedalus Implementation and Experience, S. Seshan. H. Balakrishnan, R. Katz, UC Berkeley

[37] Mobile users: To updat or not to update", A. Bar-Noy, I. Kessler, M. Sidi, Wireless Networks 1 (1995), S. 175-185

[38] An Adaptive Location Management Strategy for Mobile IP ; S. Rajagopalan, B. Badrinath, Proc. of the first ACM Mobicom 95, November 95 s

[39] IP Multicast extensions for Mobile Internetworking ; A. Acharya, A. Bakre, B. Badrinath, In IEEE INFOCOM '96, March 1996

[40] Lossless Handover for Wireless ATM, Mitts H., Hansen H., Immonen J., Veikkolainen S.; Journal of special Topics in Mobile Networks and Applications (MONET); Baltzer Science Publishers, ACM Press, Netherlands; Dec. 96

[41] Handoff and system support for Indirect TCP/IP, A. Bakre, B.R. Badrinath, 2nd Usenix Symposium on Mobile and Location-Independent Computing, April 1995

[42] I-TCP: indirect TCP for mobile hosts, Ajay Bakre and B.R. Badrinath, 15th Int'l Conf. on Distributed Computing Systems (ICDCS), May 1995

[43] A Comparison of Mechanisms for Improving TCP Performance over Wireless Links, Hari Balakrishnan, Venkata N. Padmanabhan, Srinivasan Seshan and Randy H. Katz, IEEE/ACM Transactions on Networking.

[44] Inter Access Point Protocol, IAPP Functional Specifications, Rev. 1.2; Aironet Wireless Communications, Inc., Lucent Technologies, Digital Ocean, Aug. 1996

[45] ICMP Router Discovery Messages. S. Deering. Internet Standard, RFC 1256 Sept. 1991

[46] Bootstrap Protocol (BOOTP), Croft, B., and J. Gilmore, Internet Standard RFC 951, September 1985.

[47] Dynamic Host Configuration Protocol, DHCP Internet Standard, RFC 2131 R. Droms, 3/97. Supersedes RFC 1541 and RFC 1531.

[48] Interoperation Between DHCP and BOOTP, R. Droms, Internet Standard, RFC 1534, 10/08/1993.

[49] DHCP Options and BOOTP Vendor Extensions, S. Alexander, R. Droms, Internet Standard RFC 2132 , 3/97. Supersedes RFC 1533

[50] New Router Features for IBM Network - Mixed-Media, Local Acknowledgement, LAN Manager Support, Cisco News Release, http://www-europe.cisco.com/warp/public/146/10 5.html, Jan. 1992

[51] PTOLEMY, Regents of the University of California, ftp://ptolemy.eecs.berkeley.edu, http://ptolemy.eecs.berkeley.edu, news:comp.soft-sys.ptolemy

[52] Power Saving Mechanisms in Emerging Standards for Wireless LANs: the MAC Level Perspective, Hagen Woesner, Jean-Pierre Ebert, Morten Schläger, Adam Wolisz, to appear in in Special Issue on Energy Management in Personal Communication and Mobile Computing Systems, IEEE Personal Communications System

[53] Wireless LANs: How Far? How Fast?, David Newman, Kevin Tolly, Data Communications, March 21, 1995

[54] Measuring and Reducing Energy Consumption of Network Interfaces in Hand-Held Devices, M. Stemm, P. Gauthier, D. Harada, R.H. Katz, Proc. of Mobile Multimedia Communication (MoMuC-3), Princeton, NJ, USA, Sept. 1996

[55] QUANTA: Quality of Service Architecture for Native TCP/IP over ATM networks, Sudheer Dharanikota, Kurt Maly, Old Dominion University Department of Computer Science Technical report, # TR-96-01, February 1996

[56] An Architecture and Communication Protocol for Picocellular Networks, Rohit Ghai, Suresh Singh, IEEE Personal Communications, Third Quarter 1994 , pp. 36-46.

[57] Group-based Multicast and Dynamic Membership in Wireless Networks with Incomplete Spatial Coverage, A. Bartoli, Mobile Networks and Applications Vol. 3 (1998) Nr. 2 pp. 175-188, Baltzer Science Publishers, The Netherlands

[58] Lossless Handover for Wireless ATM, H. Mitts, H. Hansen, J. Immonen, S. Veikkolainen, Mobile Networks and Applications Vol. 1 (1996) Nr. 3 pp. 299-312, Baltzer Science Publishers, The Netherlands

[59] A Buffer Overhead Minimization Method for Multicast-based Handoff in Picocellular Networks, E. Ha, Y. Choi, C. Kim; Proceedings of Mobile Communications. Technology, Tools, Applications, Authentication and Security. IFIP World Conference on Mobile Communications, Canberra, ACT, Australia, 2.-6. Sept. 1996

Glossary

This Glossary has been set up in part with the help of the IBM Software Glossary [1] and the CISCO Systems list of Internetworking Terms and Acronyms [2]

AMPS Advanced Mobile Phone Service: analogue cellular network standard, applied mostly in the USA

ARP Address Resolution Protocol: In the Internet suite of protocols, the protocol that dynamically maps an IP address to an address used by a supporting metropolitan or local area network

ATM Asynchronous Transfer Mode: International standard for cell relay in which multiple service types (such as voice, video, or data) are conveyed in fixed-length (53-byte) cells. Fixed-length cells allow cell processing to occur in hardware, thereby reducing transit delays. ATM is designed to take advantage of high-speed transmission media such as E3, SONET, and T3.

BS Base Station: A fixed device with both a radio transmitter/receiver pair and an interface to another networking infrastructure, which relays signals to and from the wired medium to and from mobile data terminals.

Beacon: repeated transmission of a beacon message used to announce the presence of a base station

Bridge: (1) A functional unit that interconnects multiple LANs (locally or remotely) that use the same logical link control protocol but that can use different medium access control protocols. A bridge forwards a frame to another bridge based on the medium access control (MAC) address. (2) Contrast with gateway and router.

Broadcast: (1) Transmission of the same data to all destinations. (2) Simultaneous transmission of data to more than one destination. (3) Contrast with multicast and unicast.

CDPD: Cellular Digital Packet Data, wireless extension service offered in AMPS networks

Cell: Coverage area of a base station, identical to its physical signal detection range

DS Distribution system: The devices, functionality and the communication infrastructure that is used to interconnect WLANs into a E-WLAN

E-WLAN Extended Wireless Local Area Network: A local area communication infrastructure consisting of several WLAN-islands and a separate interconnection infrastructure, the distribution system

Encapsulation: A technique used by layered protocols by which a layer adds control information to the protocol data unit (PDU) from the layer it supports. In this respect, the layer encapsulates the data from the supported layer. In the Internet suite of protocols, for example, a packet would contain control information from the physical layer, followed by control information from the network layer, followed by the application protocol data.

End System: (1) Generally, an end-user device on a network. (2) Nonrouting host or node in an OSI network.

GPRS: General Packet Radio Service, wireless data extension to GSM networks

GSM: Global System for Mobile Communications: digital wireless telephone network based on TDMA, originated in Europe, applied worldwide

Host: Computer system on a network. Similar to the term node except that host usually implies a computer system, whereas node generally applies to any networked system, including access servers and routers. See also node.

IETF Internet Engineering Task Force: The task force of the Internet Architecture Board (IAB) that is responsible for solving the short-term engineering needs of the Internet. The IETF consists of numerous working groups, each focused on a particular problem. Internet standards are typically developed or reviewed by individual working groups before they can become standards.

Internet Layer: In the Internet suite of protocols, the layer corresponding to the network layer in Open Systems Interconnection (OSI) architecture.

IP Internet Protocol: In the Internet suite of protocols, a connectionless protocol that routes data through a network or interconnected networks. IP acts as an intermediary between the higher protocol layers and the physical network. However, this protocol does not provide error recovery and flow control and does not guarantee the reliability of the physical network.

Internet Router: A device that forwards and routes IP datagrams from one network to another. Routers allow hosts on different networks to communicate with each other.

Internet: The worldwide collection of interconnected networks that use the Internet suite of protocols and permit public access.

LAN Local Area Network: (1) A computer network located on a user's premises within a limited geographical area. Communication within a local area network is not subject to external regulations; however, communication across the LAN boundary may be subject to some form of regulation. (2) A network in which a set of devices are connected to one another for communication and that can be connected to a larger network. (3) Contrast with metropolitan area network (MAN) and wide area network (WAN).

LLC Logical Link Control: The data link control (DLC) LAN sublayer that provides two types of DLC operation for the orderly exchange of information. The first type is connectionless service, which allows information to be sent and received without establishing a link. The LLC sublayer does not perform error recovery or flow control for connectionless service. The second type is connection-oriented service, which requires establishing a link prior to the exchange of information. Connection-oriented service provides sequenced information transfer, flow control, and error recovery.

Logical Link Control Protocol: In a local area network, the protocol that governs the exchange of transmission frames between data stations independently of how the transmission medium is shared. The LLC protocol was developed by the IEEE 802 committee and is common to all LAN standards.

MAC Medium Access Control: In LANs, the sublayer of the data link control layer that supports medium-dependent functions and uses the services of the physical layer to provide services to the logical link control (LLC) sublayer. The MAC sublayer includes the method of determining when a device has access to the transmission medium.

Medium Access Control Protocol:. In a local area network, the protocol that governs access to the transmission medium, taking into account the topological aspects of the network, in order to enable the exchange of data between data stations.

MH Mobile Host:A host which can move in a network while some property concerning it, such as its address, its connections, etc. remain unchanged

MobileIP:A set of protocols and protocol extensions developed to provide connectivity to mobile hosts in an internet environment.

Multicast: (1) Transmission of the same data to a selected group of destinations. (T) (2) A special form of broadcast in which copies of a packet are delivered to only a subset of all possible destinations. (3) Contrast with broadcast.

Node: Endpoint of a network connection or a junction common to two or more lines in a network. Nodes can be processors, controllers, or workstations. Nodes, which vary in routing and other functional capabilities, can be interconnected by links, and serve as control points in the network. Node is sometimes used generically to refer to any entity that can access a network, and is frequently used interchangeably with device. See also host.

RFC Request for Comments: In Internet communications, the document series that describes a part of the Internet suite of protocols and related experiments. All Internet standards are documented as RFCs.

Router : (1) A computer that determines the path of network traffic flow. The path selection is made from several paths based on information obtained from specific protocols, algorithms that attempt to identify the shortest or best path, and other criteria such as metrics or protocol-specific destination addresses. (2) An attaching device that connects two LAN segments, which use similar or different architectures, at the reference model network layer. (3) In OSI terminology, a function that determines a path by which an entity can be reached. (4) In TCP/IP, synonymous with gateway. (5) Contrast with bridge.

Simulcast: Synonym to Multicast

Source Routing, Source Route Bridging: In LANs, a bridging method that uses the routing information field in the IEEE 802.5 medium access control (MAC) header of a frame to determine which rings or token-ring segments the frame must transit. The routing information field is inserted into the MAC header by the source node. The information in the routing information field is derived from explorer packets generated by the source host.

Spanning Tree: In LAN contexts, the method by which bridges automatically develop a routing table and update that table in response to changing topology to ensure that there is only one route between any two LANs in the bridged network. This method prevents packet looping, where a packet returns in a circuitous route back to the sending router.

Spread Spectrum: an encoding/modulation technique for RF signals in which a signal is transmitted in a bandwidth considerably greater than the frequency content of the original information. it is particular suitable for applications where immunity to noise, interference and multipath fading is necessary. Two main principles are applied: Frequency Hopping (advantages: combating the near-far effect which is due to the non-cellular approach) and Direct Sequence (advantages: fading/jamming rejection, large address space, security).

Triangular Routing, Dogleg Routing:A routing anomaly where packets from host A to host B have to go through a third host, that must not be on the direct path between the two hosts.

Tunneling: The bypassing of normal routing procedures to get packets across a part of the network that cannot properly route them. Tunneling is usually accomplished either by source routing or by encapsulation.

WAN Wide Area Network: (1) A network that provides communication services to a geographic area larger than that served by a local area network or a metropolitan area network, and that may use or provide public communication facilities. (T) (2) A data communication network designed to serve an area of hundreds or thousands of miles; for example, public and private packet-switching networks, and national telephone networks. (3) Contrast with local area network (LAN) and metropolitan area network (MAN).

WMAC Wireless MAC Protocol: In a wireless LAN, the protocol that governs access to the wireless transmission medium in order to enable the exchange of data between stations.

Curriculum Vitae Jost Weinmiller

* 29.12.1965 in Brussels, German

Current Residence	Permanent Address:
Jan Bronnerstraat 117	Lennestr. 9
NL-2552 WX Den Haag	81234 Feldafing
Netherlands	Germany
Tel:+31-(0)70-3975663	Tel: +49-8157-1468
email: jost@xs4all.nl	Fax: +49-8157-4120

WORK EXPERIENCE

since Nov. 98	Patent Examiner at European Patent Office, Directory 1.2.49 Mobile Telecommunications, Den Haag, Netherlands
Jan.98 - Jul.98	Visiting Researcher at Cornell University, NY, USA, Wireless Networks Lab, Topic: QoS maintenance in the case of handover by applying prediction algorithms.
Nov.94 - Jul. 97	Research Assistant in DFG funded Project “MAC protocols for Wireless High-Speed LANs” Technical University Berlin, Telecommunication Networks Group, Advisor Prof. A. Wolisz Task: Evaluation and comparison of the access schemes applied in the IEEE 802.11 and HIPERLAN-1 wireless LAN standards, contributions to standardization process, evaluating applicability of the access schemes for next generation high-speed wireless LANs, presentations on national and international conferences (german and english), advising students in their master theses, administration and design of multiple WWW servers
Apr. - Nov.94	Research Assistant, Technical University Berlin, Telecommunication Networks Group, Advisor Prof. A. Wolisz Topic: Scalability issues in mobile internetworking, improvements to MobileIP
1992 - 1994	Research Assistant, Technical University Berlin, Interdepartmental Center for networking and multimedia technology (TUB-PRZ), BERKOM Project MMT (Multimedia Transportstack), Topic: Leader for Project group “ST-2” (Streams Internet Protocol), developing connection-oriented network protocol for a multimedia-capable Transportstack on multiple platforms

EDUCATION

1995 - 98	Ph.D. - Thesis, Technical University Berlin, Telecommunication Networks Group, Advisor: Prof. A. Wolisz on “Mobility Management and Interconnection of wireless cells”
1991 - 92	Master thesis written at UCLA, CA, USA, Electrical Engineering Dept, Advisor: Prof. I. Rubin, delivered at Technical University Munich, Prof. Dr.techn J. Swoboda, on “LAN - FDDI Interconnections”, Master Degree;

1988 - 91	Graduate School for Electrical Engineering and Information Sciences, Technical University Munich; Major: Information Processing;
	Relevant Courses: Computer Networks, Computer Architecture, Process Application Control, Information Technology, Speech Recognition, Satellite Technology
1986 - 88	Undergraduate School for Electric Engineering and Information Sciences, Technical University Munich

INTERNSHIPS

Oct. - Dec.90	Research Assistant at Siemens AG Berlin, Research on ESD (Electro Static Discharge) problems in automatic board assembly machines
Mar. - May 89	Research Assistant at Alcatel NV Paris, Research on differing requirements in national standards for Data Modems within the European Community
Aug. - Sept.87	Research Assistant at Magnetbahn AG Starnberg (Magnetic Trains)
	Various part-time and summer positions as technical assistant and board assembler (Siemens)

List of Authors Publications

“Evaluation of Handover Schemes for Local Area Mobility Managemen” Jost Weinmiller; MoMuC’98 - The International Workshop on Mobile Multimedia Communications, October 1998, Berlin, Germany

“Grouping Wireless Picocells with a Distribution System” Jost Weinmiller; MoMuC’97 - The International Workshop on Mobile Multimedia Communications, September 30 to October 2, 1997, Seoul, Korea, nominated for “Best Paper Award”

“Wireless LANs in Internet” Adam Wolisz, Berthold Rathke Jost Weinmiller; MWCN ‘97 First International Workshop on Mobile and Wireless Communications Networks May 12-15, 1997 Paris, France

“Interconnection of Wireless Cells - a Multicast-based Approach” Andreas Festag, Jost Weinmiller, Adam Wolisz; ICUPC ‘97 IEEE International Conference on Universal Personal Communications, San Diego, California, U.S.A. October 12 - 17, 1997

“Performance Study of Access Control in Wireless LANs - IEEE 802.11 DFWMAC and ETSI RES10 HIPERLAN” Jost Weinmiller, Morten Schlaeger, Andreas Festag, Adam Wolisz; Mobile Networks and Applications, Baltzer Science Publishers, July 1997

“Handover Protocol for Interconnecting Wireless Cells” Andreas Festag, Jost Weinmiller, Adam Wolisz; 8th IEEE Workshop on Local and Metropolitan Area Networks Berlin/Potsdam, Germany 25-28 August 1996

“Wireless Access to High Speed Networks” A.Wolisz, M. Schlaeger, J. Weinmiller, H. Woesner; High-Speed Networking for Multimedia Applications, W. Effelsberg, O. Spaniol, A. Danthine, D. Ferrari (eds.), Kluwer Academic Publishers, Boston/Dordrecht/London, 1996

“Analyzing and Tuning the Distributed Coordination Function in the IEEE 802.11 DFWMAC Draft Standard” Jost Weinmiller, Hagen Woesner, Adam Wolisz; June 1995, MASCOT ‘96 - International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunication Systems; Feb. 96, San Jose, California

“Modified Backoff Algorithms for DFWMAC‘s Distributed Coordination Function” Jost Weinmiller, Hagen Woesner, Jean-Pierre Ebert, Adam Wolisz; 2. ITG Fachtagung Mobile Kommunikation ‘95, Sept 95, Neu-Ulm, Germany

“Analyzing the RTS/CTS Mechanism in the DFWMAC Media Access Protocol for Wireless LANs” Jost Weinmiller, Hagen Woesner, Jean-Pierre Ebert, Adam Wolisz; IFIP TC6 Workshop Personal Wireless Communications (Wireless Local Access) April ‘95 Prague, Czech Republic